FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Polypharmacy and falls in the elderly population

Tiago Adelino Araújo Rocha

Master in Informatics and Computing Engineering

Supervisor: Henrique Lopes Cardoso

21st January, 2013

Polypharmacy and falls in the elderly population

Tiago Adelino Araújo Rocha

Master in Informatics and Computing Engineering

21st January, 2013

Abstract

The syndrome of falls is a common and serious problem and if we are considering the older population this problems takes even greater proportions. Falls can today be associated with multiple causes, which are related with factors that can be hard to track. In order to prevent side effects, among them falls, from happening medications usually lists possible unwanted symptoms. However, when a patient is using multiple drugs side effects may happen due to the combination of medicine.

In order to identify combinations and single pharmacological drugs that may have some side effects related to falls, new community based solutions must be developed. In this project we developed a system capable of analysing and evaluating the fall risks associated with medical drugs usage on the elders. This system must be capable of making this evaluation regarding a single drug or a drug combination, pointing out the riskier drugs or combinations.

We used a process that classifies the drugs according to their classes, types or families and assess if that classification represents a group or a specific substance of risk to this scenario. The system was capable of making the evaluation within reasonable time and provide the analysis conclusion to the user. The system achieved the expected results with a success ratio of 90% for single drug evaluations in a test with 93 drugs and 70% for drug-drug interactions in the conducted tests, the system has proven to achieve the expected results.

Resumo

A ocorrência de quedas é um problema sério e comum nos dias de hoje. Se considerarmos unicamente a população mais idosa este problema toma proporções ainda mais alarmantes. O acontecimento de quedas está actualmente associado a várias causas que estão ligadas a factores que podem ser dificeis de determinar. Para prevenir consequências relativas aos efeitos secundários, a medicação actualmente contem um panfeto que lista os possiveis efeitos adversos das substâncias a tomar. No entanto, quando um paciente utiliza múltiplos medicamentos podem surgir efeitos secundário adversos provenientes dessas combinações farmacológica.

De forma a identificar combinações medicamentosas e medicamentos unicos que possam ter efeitos secundários relacionados com quedas, novas soluções e estudos baseados em comunidades tem de ser desenvolvidos. Neste projeto temos como objectivo desenvolver um sistema capaz de analisar e avaliar o risco de queda associado ao consumo de medicamentos nos idosos. O sistema deve ser capaz de analisar a utilização de medicamentos separadamente e em conjunto, tendo a capacidade de detectar efeitos provenientes das combinações.

Durante o desenvolvimento do projeto utilizamos critérios que classificam os medicamentos de acordo com a sua classe/tipo/familía e a partir desta informação verifica se esse medicamento faz parte de um grupo potencialmente perigoso. O sistema final revelou em testes ser capaz de avaliar correctamente e rapidamente 90% dos medicamentos que lhe foram apresentados. Para avaliação de combinações medicamentosas o resultado atingido foi de 70%, atingindo assim a fiabilidade esperada.

Acknowledgements

First and foremost, i would like to thank my supervisors of this project, Dr. Liliana Ferreira and Prof. Henrique Lopes Cardoso for valuable guidance, suggestions and advice through all the steps of this thesis development.

To Fraunhofer AICOS Portugal for the awesome opportunity, providing the necessary means and interaction with skilled professionals.

I would also like to specially thank my parents and family for the continuous support and understanding provided during my academic training through all this years.

To my high school colleges, today great friends, Filipe, Ana, André, Machado, among others, for the good moments and support during all this years.

Not forgetting the countless nights working or parting, to all my college friends for the help in those panic moments before the exams or deadlines.

Finally to my cat Branca and my dogs, Max and Snoopy!

Tiago Rocha

"A dream doesn't become reality through magic;
it takes sweat, determination and hard work."

Colin Powell

Contents

List of Figures

LIST OF FIGURES

List of Tables

Abbreviations

FhP	Fraunhofer Portugal
WHO	World Health Organization
ICECI	International Classification of External Causes of Injury
TCAs	Tricyclic Antidepressants
SSRIs	Selective Serotonin Reuptake Inhibitors
PHR	Personal Health Record
OWL	Web Ontology Language
UIMA	Unstructured Information Management Architecture
AICOS	Assertive Information and Communication Solutions
W3C	World Wide Web Consortium
XML	Extensible Markup Language
UMLS	Unified Medical Language System
SPARQL	SPARQL Protocol and RDF Query Language
OWL	Web Ontology Language
RDF	Resource Description Framework
PDF	Portable Document Format
CNS	Central Nervous System
BZD	Benzodiazepines
TCA	TriCyclic Antidepressant
SSRI	Selective Serotonin Reuptake Inhibitor
SNRI	Serotonin-norepinephrine reuptake inhibitors
ICD-9-CM	The International Classification of Diseases, Ninth Revision, Clinical Modification
HCUP	Health Cost and Utilization Project
ATC	Anatomical Therapeutical Classification System
PHARMINX	PHARMacological INformation EXtraction System
UTS	UMLS Terminology Systems
NCI	National Cancer Institute
MEDRA	Multilingual European Drug Registration Agency
MESH	MEdical Subject Headings
ICPC	International Classification of Primary Care
WHOART	World Health Organization Adverse Reactions Terminology
CUI	Concept Unique Identifier
SIDER	SIDe Effect Resource
SFINX	Swedish, Finnish, INteraction X-referencing database

ABBREVIATIONS

EMR	Electronic Medical Record
API	Application Programming Interface
DIC	Drug Interactions Checker
DDI	Drug-Drug Interactions
SDEP	Single Drug Evaluation Process
LD	Levenshtein Distance
NED	Normalized Edit Distance
AAFP	American Academy of Family Physician

Chapter 1

Introduction

Nowadays, the general profile of the population in most countries is a lot different from what it was 20 years ago. With the technological and scientific evolution in the last decade the health care service has improved exponentially. All these factors together have resulted in the ageing of population and an increase in number of elderly population.

Based on data collected from [WHO13], the life expectancy specifically in Portugal has increased from 74 years on 1990 to 79.4 years on 2009, which means an increase of 5.4 years, a significant increase even for a developed country. It is possible to notice a clear discrepancy between the developed countries life expectancy and 3rd world countries in the available data. Africa's life expectancy in 2009 is about 54 years old, in contrast to the Europe and America's which is 20+ years higher.

The elderly population will have a higher falling rate due to age-related changes and frailty, so, falling can eventually lead to very serious physical and/or psychological outcomes. In the worst case scenario it can lead to death, but it can also lead to incapacitating injuries, depression, fear, suffering, loss of independence and as a result, decreases in life quality. Those kind of scenarios can become a problem that will expand to the family of the patient, since it will have economical and social consequences and will be time consuming[Ga11].

The recent increase in numbers of the elderly people places the problem of falling in the community as one important matter to be attended, since the health care keeps improving and there will still be an increasing portion of the population in elderly stage of life. We will reach a point where it will be too difficult to keep track of all the registered cases and accompany them individually in a non automated way.

It is now proven by a vast number of studies conducted all over the world[KPY+03, HLL07, RCV03, EBM02, ZDH+06] that fall prediction by multi-medication is an issue and can be prevented or at least minimized by changing some patient habits and adopt precautionary measures in their daily activities[Dye07]. An article on American Academy of Family Physician (AAFP) states that:

> *The side effects of some medicines can upset your balance and make you fall. Medicines for depression, sleep problems and high blood pressure often cause falls. Some medicines for diabetes and heart conditions can also make you unsteady on your feet.* [Phy00], p.2173

A better watch on the patient during a critical stage of the medication will decrease the risk of falling and consequently, the effects related to the fall. If we could be capable of identifying that critical stage the advantages would be immense.

Currently, are several studies under development that research on the side effects of multiple medications. However, the majority of those studies are based on a fixed population limiting the scope of results to the events verified in that scenario. In this regard, this project aimed at developing a tool capable of analysing the medication being used by a patient and assess the situation by identifying the risks related to falls in a specific given scenario regarding the amount and types of medication. This will ideally help to prevent possible falls and consequently injuries or traumatic experiences that could lead to a decrease in quality of life for the patient and their families.

This dissertation intends to study the relation between the single, multi-medication and the falls in the older population. A specific tool was developed for this purpose. We have identified types, families and groups of drugs in order to identify their substances and possible combinations that might lead to a fall risk increase.

1.1 The Problem

Falls are a serious concern nowadays, specially in the older population. There are some methods, studies, initiatives and tools available in this field to prevent, predict and identify falls [FPC05, Agi10, SPA09], but they revealed themselves not viable to this specific purpose of risk identification on drug usage, mainly because it is not their specific goal. With the recent increase of the older population this problem became even more serious. Therefore fall prevention is becoming a challenge of greater importance to the well being of the elder population.

It is a proven fact that medication can lead to a higher risk of falling [Phy00, KPY+03, HLL07]. The side effects of each drug are available on the leaflets, this includes all types of associated side effects and may even mention fall risk itself. However, with the combination of 2 or more drugs, unreported and unexpected side effects may occur. The usage of 2 distinct drugs with individual side effects unrelated to falls can lead to a side effect that will increase the falling ratio. Some of the combined side effects are already reported and known, but due to the huge number of medicine available to us, it is currently impossible to report all the combined side effects. The known side effects from the drug interactions are normally checked by the healthcare professional when prescribing multiple drugs, but since it is a manual process there may be some details or drug interactions that may go unnoticed.

The majority of existing studies regarding this matter are performed in a clinical setting and are only available to a limited part of the population. In this project we intend to use some data

gathered by these studies to identify the main drug families/groups/types that are related to high fall ratios.

The technological development and computational power available today will make it possible for us to set up an automated system capable of analysing a patient's drug usage and identify the fall risk according to the drugs properties and interactions.

1.2 Project

Polypharmacy effects in the elderly population is a dissertation that was proposed by Fraunhofer Portugal Research Center for Assistive Information and Communication Solutions (Fraunhofer AICOS)[1]. Currently this institution is developing several projects related to falls, mainly with the objective of using the sensors available in mobile phones to evaluate the ambulatory information. By recording and tracking the way people move it is possible to study the occurrence of falls and assess if the user is standing, sitting, ruining, among other possible states.

It is possible to apply this knowledge in a vast number of subjects, and getting the record of the medication currently used by them will allow to make an analysis on the side effects of the polypharmacy in multiple cases.

Our specific part in this, is to develop a tool to evaluate the relation between the medication being taken by a patient and the fall risk associated with that same medication. This work could eventually be integrated with the currently existing projects in Fraunhofer AICOS to create a more complete and effective evaluation/prediction tool for falls.

In the development of the project a lot of information about the individual drugs that are being used in different cases will be available. We will identify the drug, link the drug to a specific concept, and from there evaluate that drug according to the available databases. Knowing the type, family or even components of the drug makes it possible for us to make a study on the correlated fall risk that is associated with that same drug. The information gathered will be stored in a pre-developed and pre-populated knowledge representation model that we will study and improve with new structures and information related to our study.

1.3 Goals and Motivation

Since currently the portion of the elderly population is increasing it is of great importance to provide a system that can do something about the vast array of problems related to multi medication and falls that are presented in this part of the population. The fall itself represents only a part of the problem that we intend to solve, or at least minimize. It is usual nowadays to see older people using multiple medication daily, sometimes for the same disease, and the negative consequences of the mixed side effects can, in some cases be very severe. Those cases that may happen without notice from the healthcare professional and may bring severe consequences to the patient.

The project has the following main objectives:

[1] http://www.fraunhofer.pt/en/fraunhofer_aicos/about_us.html

- improve/update a knowledge representation model that will be capable of representing the medical drugs and relevant associations
- Identify the system entries as substances or trade names
- If the entry is a trade name the system must be able to identify the active substance
- Relate the entries to the respective concept or create the concept if not found
- Update the concept if possible with external data regarding the drug class/family/type
- Update the side effects if possible/needed
- Evaluate the risk of the entry drugs individually according to the gathered information
- Evaluate the drug interaction risks to the given drug set

Doing all the steps of this process using human resources would take a lot of manpower and research time.

The final goal of this project is to make possible the analysis of the usage of multi-medication in the elderly populations automatically. This project will focus exclusively on the components and side effects related to the falls.

1.4 Structure of the document

This document is divided in 7 different chapters.

Chapter 2, background and literature review, reports the study of the problem and explains the reasons for the problems as well as the possible ways to solve them as well as the existing projects related to the thesis subject.

Chapter 3 specifies the platforms and necessary information for this project. Here we conduct a detailed study of the technologies we used and sources for the information we need to proceed to the drug evaluation.

Chapter 4 specifies in detail the modification performed in the existing data model and structure in order to adapt it to our purpose. We will justify the modification individually and explain the approach.

Chapter 5 explains how the core of the system works. It will explain all the process and data access used during the fall risk evaluation process. The whole process and sub processes are exposed using diagrams and textual explanation.

Chapter 6 contains information on the matching tests and verification of the accuracy of the various results we've reached following the approach presented in chapter 5.

Finally, conclusions are presented in Chapter 7 where we mention the problems we have overcome and possible directions to address the ones that persist.

Chapter 2

Background and Literature Review

This chapter will present a literature review on polypharmacy, possible solutions and important background concepts. This review will be focused in the elder population. The information will be presented considering the project description and the proposed objectives, so that the proposed approach can be introduced upon a careful investigation of the background existing literature related to falls and polypharmacy in the elderly population.

2.1 The ageing process

In the last century, the average life expectancy has increased substantially, mainly in the developed countries [WHO13] . This happens due to a series of changes in lifestyle, personal habits, sanitation, science and technology in the society. All these factors will contribute to decrease disease-related deaths in the population, and consequently people will have a longer life expectancy which will lead to an increase in the elderly population [FTW+01].

The figure 2.1 represents the life expectancy per country in Europe. The information available dates from 2009 and the source is the World Health Organization [WHO13] .

The study of falls in this thesis will focus on the elderly people. "Elderly people" is a vague definition, because it depends on the place, culture and economics where we apply it, but in absolute terms and for this thesis "Elderly people" is considered to be someone 65 years old or more.

Since birth, a human being's body will be constantly accumulating changes which will gradually increase susceptibility to disease and death. This natural process that occurs in every living being is called "Ageing process" and its definition according to [Har81] is the following:

> *Ageing is the progressive accumulation of changes with the time that are associated with or responsible for the ever-increasing susceptibility to disease and death which accompanies advancing age.* [Har81], p.7125

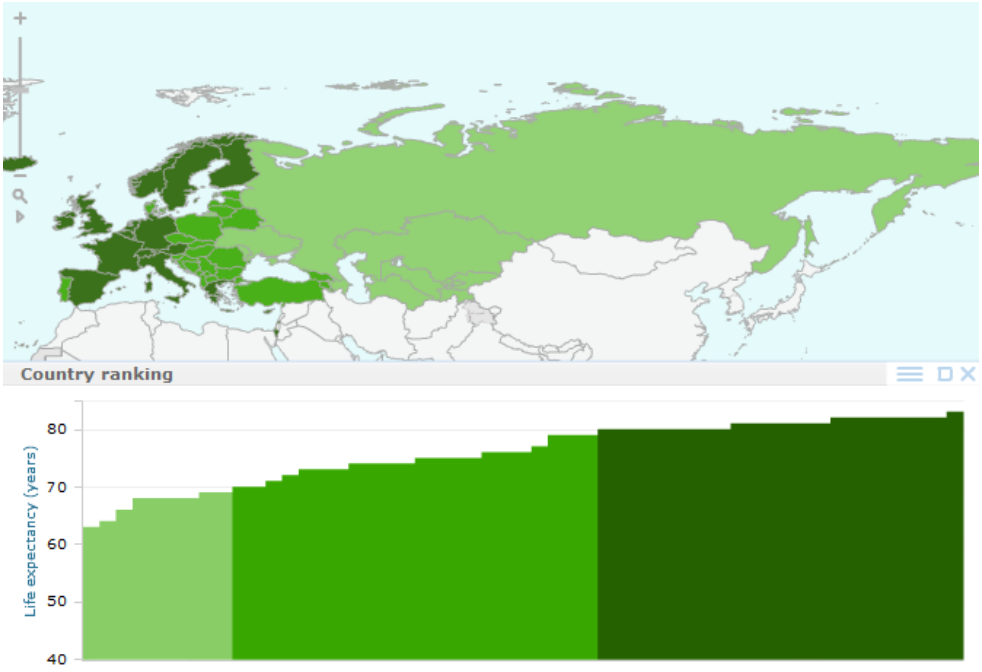


Figure 2.1: Life expectancy in Europe 2009

The wear of the organs and body will depend on the care and usage during the patient's life. Due to this "Ageing process" the older people will consequently have a lower capability to deal with eventual diseases, virus, falls or other kind of health issues. They will also eventually start to lose cognitive and physical skills that are as well related to the ageing process.

The ageing itself is not the only process that contributes to a decreasing in the cognitive and physical skills. This is influenced as well by genetics, environment and social factors. With all these variables and in such a vast population we will have a lot of subject profiles with the same age that may behave differently with the variation of the factors mentioned.

The ageing process can be separated in 3 distinct modes. First we have the usual/normal ageing process which refers to the most common way people age. This first process is defined by a gradual decline in physical skills which will lead to other problems like diseases, loss of mobility and other problems that will eventually result in complications and death. Then we have the pathological ageing. Pathological ageing is different from usual ageing since it is applied to a group of subjects that are predisposed to certain diseases or have negative lifestyle habits that may have serious consequences leading to a premature death.

The other ageing model is the successful ageing, which may be mistakenly interpreted as a longer life, but instead is a qualitative reference to the ageing process where the usual ageing problems do not manifest so intensely and the subject can have a life quality above average for his/her age. This kind of ageing can be achieved by regular physical care, a balanced diet and avoiding bad habits like tobacco or drugs [Gar10].

2.2 Falls

Falls can be normally associated with old age. The fall itself may not be the main problem in this case, but the long or short term injuries caused as a consequence of the fall and the wear of the body. This is a problem that will aggravate according to the age and other factors of the victim. Patients with diabetes will have trouble healing possible wounds and a huge array of complications may come from there. Due to the increasing numbers of elderly people within our community the study of the falls in older people has been increasingly important in order to prevent long and short term injuries and improve the quality of life in this part of the population.

Definition

There are many causes to a fall and many definitions of falls according to the cause or motivation of the fall itself. Here we will check some literature and medical concepts on falls.

According to the International Classification of External Causes of Injury (ICECI)[ICE04], a international entity related to the World Health Association (WHO) that is responsible for the classification of all the external causes of injury, a fall is a descent or drop by force of gravity, *"nonsyncopal event not attributable to sustaining a violent blow, loss of consciousness, stroke or epileptic seizure" [ICE04].*

A book published by the Agency for Healthcare Research and Quality [SH08] states that a fall may be precipitated by intrinsic or extrinsic factors. Intrinsic factors are the ones that have a physiologic origin, and extrinsic factors are those precipitating from environmental or other hazards. According to that same book, the definition of fall is

> *"an event which results in a person coming to rest unintentionally on the ground or lower level, not as a result of a major intrinsic event (such as a stroke) or overwhelming hazard" [SH08].*

The International Classification of Diseases, Ninth Revision, Clinical Modification, (ICD-9-CM) [1] produced a series of codes that are used to categorize and separate the various types of falls. There we can see the following description:

> *Accidentally bumping against moving object caused by crowd with subsequent fall (E917.6); Fall on or from ladders or scaffolding (E881); Fall from or out of building or other structure (E882); Other fall from one level to another (E884); Fall on same level from slipping, tripping, or stumbling (E885); Fall on same level from collision, pushing, or shoving by or with another person (E886); and Other and unspecified fall (E888).48 In the inpatient care setting, E888 is the code that is typically used to record a fall in a medical record.*

Morse [Mor09] says that falls can be classified in 3 different categories:

[1] http://www.cdc.gov/nchs/icd/icd9cm.htm

- **Accidental falls** - related to extrinsic factors, such as environmental factors like slipping, tripping or some other kind of mishap. Other causes for this kind of fall are errors of judgement, such as leaning against a curtain thinking it was a supportive wall. Accidental falls may also occur if the patient loses balance when ambulating.

- **Anticipated physiologic falls** - caused by intrinsic physiological factors, such as confusion os other kind of disorientation. The kind of falls that normally occur with the patients identified as fall-prone due to the high falling risk.

- **Unanticipated physiologic falls** - derived unexpected intrinsic events, such as a new onset syncopal event or a major intrinsic event such as stroke. Basically this kind of falls will occur in patients that apparently are not likely to suffer from a fall due to physiological factors. It is the kind of fall that cannot be predicted before the first occurrence.

2.2.1 Incidence

Here we will illustrate and study that concept based on real studies and data, starting with a technical report by the public health Agency of Canada [SPP10] . This report is based on the death statistics since 1997 to 2002. The report only uses data from people with 66+ years old and focused on deaths due to unintentional falls.

Figure 2.2 represents the distribution of fall occurrences in a limited sample of the population and divided by age group [SPP10]. It is clear from an analysis of the graph that the age is a factor on the occurrence of the falls.

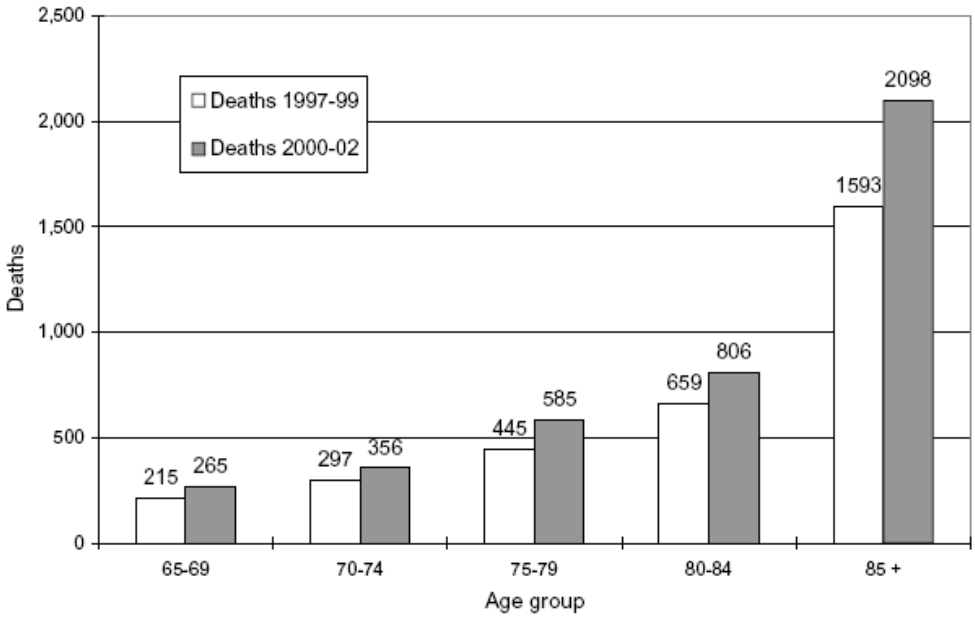


Figure 2.2: Deaths due to falls, by age group, age 65+, Canada, 1997-2002 [ORM09]

Another research [ORM09], by the Healthcare Cost and Utilization Project (HCUP) reached conclusions regarding the evolution of the falling patterns with age. Figure 2.3 shows in a graph the recurrence of falls according to the age of the subjects.

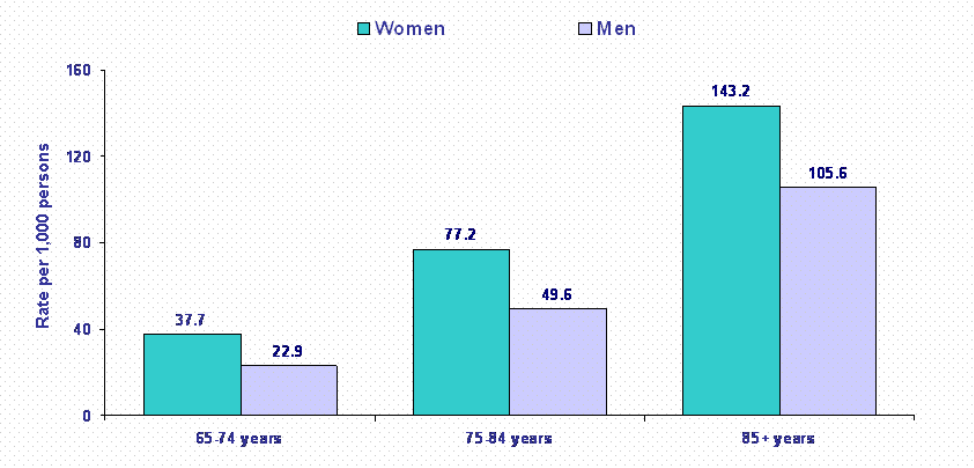


Figure 2.3: Rates of Emergency Department visits for injurious falls among elderly

2.2.2 Impact and Consequences

According to Tonarelli [Ton10] falls occur every year in 1/3 of the population with >65 years old. This statistic will take even larger dimensions if we talk about those who already had a fall. Here we will study some of the possible outcomes of a fall and the impact that may be forced upon the victims of falls.

- **Hip Injuries** - the result of falling in the elderly result most of the times in bruises, traumas and fractures, mainly on the hips, spine, arms, ankles or legs. Reports from the American Academy of Orthopedic Surgeon say that approximately 1/4 of those who have this injury die within a year and 50% lose mobility. This means that the victims of this type of injury will become dependent on their daily life which may cause other type of problems. Hip injuries are a rare problem in the younger people, but it is a kind of injury whose probability increases with age. For example, the rating on hip injuries on 85 years old subjects are 15 time higher compared to that same problem on a 56 years old patient [Ton10].

- **Impaired Functioning and Mobility** - Hip injuries are only one of the possible consequences of a fall. We can name some other grave injuries that can change drastically the daily life of a person. It is normal for everyone to gradually lose some function of their bodies over the years and that can lead in some cases to the lost of independence. Simple things to a middle age person like climb stairs, dress, rise from a chair or bed, walk and others things can no longer be doable to the elderly due to the age alone. A study published in the journal Age and Ageing [MGMO12] revealed that functioning decline was observed in

more than 35 percent of elderly who had fallen at least once in the previous year. 15% of the studied population had more serious mobility problems with basic functions like walking, cycling and other kinds of basic mobility functions. The main sub consequence of mobility lost is that it is a exponential problem. The more mobility an elder looses the more it will loose in the future due to the lack of body usage and degradation of the muscular tissues and bones.

- **Fear of Falling** - the same study mentioned above reached a conclusion that shows that falls can have a social consequence on the patient. This means that the victim of the fall can withdraw from social activities like visiting friends or relatives or even go out for a walk. The mobility problems that can result from the fall may have an important part in this, but the fear will contribute in part to the decrease of basic activities. The lost of confidence in the ability to move around is the reason for the development of what we mention here as "fear of falling". The study from Vrije University shows why the victims of falling tend to isolate themselves and all the consequences mentioned above may cause vulnerability, depression and another psychological problems that will inevitably reflect in the quality of life. According to [SSvD+08] fear of falling will lead to similar consequences as the ones presented to impaired mobility function:

  > *The main consequences were identified as a decline in physical and mental performance, an increased risk of falling and progressive loss of health-related quality of life.*

2.3 Side Effects

The likelihood of illness is increased with the ageing process. Directly connected to this logic is an increase in medical drugs usage. A study from the year 2000 [MB00] reveals that in the USA people over 65 years old consume over 40% of the sold medicine, even though they only represent 13% of the population. Today the life expectancy is slightly higher, but the tendency is the same. The same scenario is presented to us in the UK where 65+ years old populations represent 18% of the total population but still consume 45% of the medical supply in the country [MB00]. The normalization on the usage of such high number of medical drugs will cause for some side effects to go unnoticed. If we add the low health expectations on this kind of patients that may make them less likely to complain we have a vast set of factors that contribute to the side effects accumulation. This is a problem that may be solved with some work by the health professionals.

An extensive search revealed that a list of side effects related to falls is not available, at least a complete and validated one. In this case scenario the obvious choice was to investigate and create one. One way for us to check for the fall risk of a specific drug is to evaluate the side effects directly related to that medicine. In order to do this it is crucial for us to do a background check on fall related side effects. Another way for us to investigate the most common side effects related

to falls is to evaluate the recurrence of side effects in potentially risky drugs. So, for us to get the maximum number of fall related side effects we will use both methods and present the results.

Background

According to a study [RJR94] the most usual side effects connected to fall include gait and balance disorders, weakness, dizziness, environmental hazards, confusion, visual imparity and postural hypotension, lower-extremity weakness, gait and balance instability, poor vision, cognitive and functional impairment. In this list, some connections to falls may be obvious, but others not so much. In this study, the most common problems were related to gait and weakness problems that result in nearly 1/4 of the fall cases. Lower extremity weakness was another big contributor to the fall ratio, as 48 to 57% of the patients suffer from this. Some of these problems exist on this kind of population only due to the fact of their age. The drug side effects normally act as an aggravating factor that will increase the severity of those problems.

According to [MGMS80], healthy older persons will score on average 20 to 40% lower on strength tests when compared to young people. We still have to take into account that a good part of institutionalized older people are not healthy, so the combination of all these factors will increase exponentially the related risk.

Awareness or mental disparity effects like poor vision, cognitive impairment, confusion and dizziness can easily be associated with falls since the subject will lose motor, evaluation and reacting time capabilities. Another type of issues that may lead to falls are those related to physical problems or disabilities. Side effects like gait balance, weakness, lower-extremity weakness, balance instability and cognitive impairments are listed on this class of problems due to the difficulty in reacting or avoiding a fall situation by the subject.

2.4 Polypharmacy

The motive that leads to a fall can vary and we can identify a series of risk factors. One of the main risk factors in older people is correlated with certain types of medication, and the combination of multi-medication generally increases the associated risk. This case scenario is even more dangerous when the mixed drugs are already by themselves considered of high risk for fall related problems.

Here we will discuss the effects of various types of medications on falls so we can evaluate the associated risk per medication type. Some investigations/studies done before reached conclusive results between the medication types and the probability of falling in older people.

Most of the studies we consulted were done targeting a specific part of the population. The limitations of each study vary on the gender, age group and targeted drug family/group/type. The vast majority of the studies are done in a controlled trial and then, achieve conclusions when compared to a control group. Some other studies use records from healthcare institutions and limit them by the trial by specific requirements. A study published in the British Journal of Clinical Pharmacology [ZDH+06] revealed that it is possible to assess a fall risk based only in the number

of drugs that the subject is currently taking. However, this kind of risk measurement is not accurate and may be misleading, since it does not take into account many other fundamental factors. Figure 2.4 represent the results by the British study [ZDH+06]:

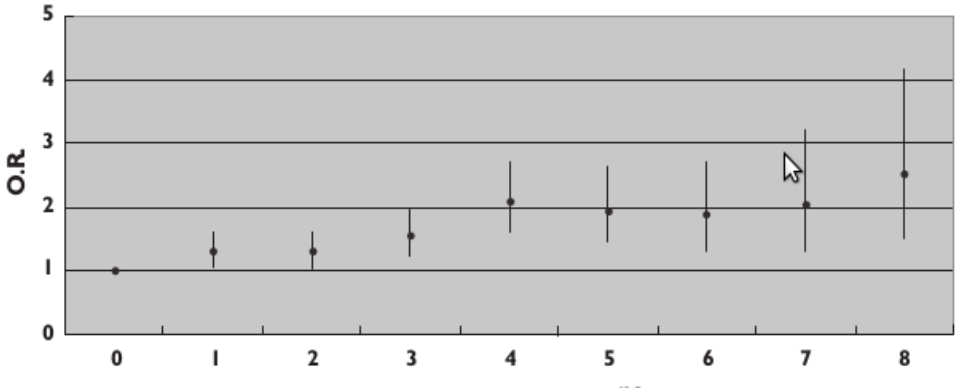


Figure 2.4: Polypharmacy fall risk according to the number of drugs[ZDH+06].

We can clearly see a direct relation between the number of drugs taken (X axis) and the Fall Odd Ratio(Y axis) value in comparison to a control group, which proves that there is a relation as we anticipated. The study also states that with the increase in the number of drugs being taken, the probability of taking risky drugs also increase. This is why the verification of the drug types and risk must be done for all drugs being used.

In the study, the drugs considered risky were mainly Central Nervous System drugs and diuretics. For us to get definitive conclusions we have to evaluate the results of multiple studies/clinical trials.

2.4.1 Definition

The term polypharmacy refers to the group of medications one person may be taking. It comes from two Greek root words: *poly*, meaning many, and *pharmakeia* meaning medicines or drugs. It is generally used when that one person is taking multiple medications, or when the drugs have been prescribed by many doctors, and may not have been coordinated well. This is usually a problem for adults with advanced age that may have more than one medical condition. Since they need medication to more than one disease there could be some extra symptoms due to the mix of medication, side effects that are not mentioned in either one of the original medicine.

2.4.2 Definitions and classification of medication types

At this point it is important to say that medication types, families and groups overlap very often. A family of drugs can contain substances from various families or a substance may be a part of several drug groups. We will mention the cases we have found during this literature review.

2.4.2.1 Benzodiazepines

Several other studies [PMT+01, PVR+00, NHMP96] revealed benzodiazepines as one of the most dangerous drugs when we are talking about falls in the elderly. Benzodiazepines are a chemical group whose chemical core structure is composed by benzene and diazepine. Benzodiazepines may be a subgroup of the Antidepressant ,Hypnotic, Sedative, Anxyolitics, Antiepileptics, and some other Central Nervous System(CNS) related drugs. During this study we found out that this kind of drugs have a top priority when it comes to fall risk prediction.

The figure 2.5 represents a part of the study [PVR+00] where we can clearly see the impact of benzodiazepines, among other factors, on the patients. The first column represent the Odd Ratio when compared to a control group and the second column represents the standard deviation.

	Odds ratio (95% confidence interval) and P value		
Age groups (years)			
<65	1		
65–80 years	1.6	(0.9–2.8)	0.07
>80 years	2.7	(1.5–4.7)	<0.001
Benzodiazepines			
Nonuser	1		
half-life <6 hours	1.9	(1.03–3.3)	<0.05
half-life 12–24 hours	1.8	(1.2–2.8)	<0.01
half-life >24 hours	0.8	(0.4–1.8)	n.s.
Combination of BDZ	1.6	(0.8–3.3)	n.s.
Other psychotropic agents	2.3	(1.6–3.2)	<0.001
Antidiabetic agents	1.5	(1.03–2.2)	<0.05
Number of drugs			
0–2	1		
3–5	1.2	(0.8–2)	n.s.
>5	1.6	(1.02–2.6)	<0.05
Charlson Index			
0	1		
1–2	1.01	(0.7–1.6)	n.s.
⩾3	1.7	(1.05–2.8)	<0.05
Cognitive impairment (%)	1.6	(1.08–2.3)	<0.05
Length of recovery (days)			
<10	1		
10–16	1.6	(0.99–2.5)	0.054
⩾17	2.1	(1.4–3.3)	<0.001

Figure 2.5: Multivariate analysis of risk factors for falls in 7908 hospitalized patients[PVR+00].

2.4.2.2 Antidepressants

Antidepressants are a psychiatric medication which is normally used to alleviate mood disorders, such as major depression and dysthymia and anxiety disorders such as social anxiety disorder. The effects of antidepressants will take time to take action on the subject and are usually administered for long periods of time, from months to years.

There are two main types of antidepressants, the selective serotonin reuptake inhibitors (SSRIs) and the tricyclic antidepressants(TCAs). Generally, in older patients the use of TCAs is preferable; however, SSRIs are associated with a reduced fall ratio [RBBL03] .

On the other hand, another study revealed that the fall ratio was slightly higher with the use of SSRIs (6.3%) compared to TCAs (4.76%) [HLL07] . Regardless, both studies agreed that the risk of falls remained high in the long-term usage of antidepressants. The worst case scenario is verified when a patient uses both types of antidepressants which result in an even higher fall ratio. This ratio is directly proportional to the antidepressant dosage.

TCAs and SSRIs were unidentified in various articles or clinical trials and proven to be one of the main drugs classes related to falls in older persons [Hub03, AWA01, LAM+98].

> *With participants who had no exposure to antidepressants as the reference category, the adjusted odds ratio for hip fracture was 2.4 (95% CI 2.0–2.7) for exposure to SSRIs, 2.2 (1.8–2.8) for exposure to secondary-amine TCAs, and 1.5 (1.3–1.7) for exposure to tertiary-amine TCAs. For all types of antidepressants, current use was associated with a higher risk of hip fracture than former use. The odds ratios for hip fracture were higher for new current users than for continuous current users in all three drug classes. The proportion of current use in the low-dose range was 22% for SSRIs, 50% for secondary-amine TCAs, and 58% for tertiary-amine TCAs.* [LAM+98], p. 1303

2.4.2.3 Antihypertensive drugs

Antihypertensive drugs are a type of drugs used to treat high blood pressure or hypertension in order to avoid complications like strokes or myocardial infarction. The blood pressure-lowering effects can result in some side effects like orthostatic hypotension and dizziness that represent a serious problem.

Orthostatic hypertension is a form of hypertension in which a person's blood pressure suddenly falls when standing up or stretching and can occur briefly and occasionally in anyone, but it is particularly usual among the older people.

With this kind of drug in the right dosage it is possible to improve the cerebral and systemic vascular response to hypotensive stress. Consequently, there is no association of antihypertenside drugs with an increased risk of falls. So, and according to [RBBL03] , there should be no reason for a major falling ratio problem associated with Antihypertensive medicine.

2.4.2.4 Anticonvulsants/Antiepileptic drugs

Anticonvulsants, also known as Antiepileptic drugs are prescribed to treat a range of medical conditions like seizure disorders and epilepsy. Among those are diseases like bipolar disorder, fibromyalgia, and pain caused by nerve injuries or other conditions. There are reports that associate this kind of drugs to be directly linked to an significantly higher risk of falling in elderly people, specially women. The effects of anticonvulsants in the central nervous system was identified as the cause for this behaviour [RBBL03]. It is important to say that anticonvulsants may contain a part of other risky drug classes like Beznodiazepines.

2.4.2.5 Anxiolytics

Anxiolitics are another CNS drug. Several studies and clinical trials [LPE03, MM97, RCV03] revealed Anxiolitics and other CNS drugs to have a direct relation with the fall ratio verified in older and some middle age subjects. There are some anxyolitic drugs that are benzodiazepines. Those are the main cases to worry about and where we should focus our attention. The remaining anxyolitics also present some increase in fall ratio but with much less significance. Anxiolytic drugs are also known as minor tranquillizers.

2.4.2.6 Beta Blockers

Beta blockers are a class of drugs that, as the name states, block the beta receptor agents. This kind of agents are part of the sympathetic nervous system and are responsible for actions related to stress reaction and adrenaline situations. The normal target for beta blockers are problems related to cardiac arrhythmias and hypotension. The relations of beta blockers and falls is studied in [RCV03] where we can read that:" *Recurrent falls were reported 2.0 times as often among beta-blocker users as among non users, after adjusting for cardiovascular disease*".

2.4.2.7 Muscle Relaxants

As the name indicates, muscle relaxants are all the drugs that affects skeletal muscle function in order to decrease the muscle tone. Normally the usage of this type of drugs is to decrease involuntary movements or spasms. The connection to falls here is pretty straight forward. These effects may reduce the capacity of the subject to react in a fall situation by reducing the response time of the muscles or even cause that same fall situation.

2.4.2.8 Opioids

Opioids are also refereed in some articles as a possible fall contributor. However, this association is controversial, since while there are studies that revealed a distinct and clear relation of these drug types to falls [EBM02], other studies says that there are no indicators to reach that conclusion [KPY+03].

2.4.2.9 Psychotropics

An article in the Journal of gerontology [HLL07] reviews multiple clinical trials ad combines all the available information on medication and falls in order to reach some interesting conclusions. This study says that:

> *Central nervous system drugs, especially psychotropics, seem to be associated with an increased risk of falls. The quality of observational studies needs to be improved, for many appear to lack even a clear definition of a fall, target medicines, or prospective follow-up. Many drugs commonly used by older persons are not systematically studied as risk factors for falls.*[HLL07], p.1

Almost all of the drug types listed above are Central Nervous System (CNS) drugs that are known to have some unwanted side effects, mainly related to the individual perception and body control. After checking individually the medicine types listed above it is important to focus on the fact that the conclusions pointed for each drug type individually may not be the same if we combine two or more of them. When the scenario is different and the patient is taking several types of pharmacological drugs it is possible to observe different outcomes that are not listed in any of the medications.

In order to avoid undesirable side effects related to falls it is possible to use medication in order to to cancel, minimize or maximize those side effects. Sadly, this solution is not viable due to the costs and possible complications related to the side effects of polypharmacy. We can mention as another possible scenario, we may consider, the usage of 2 drugs, neither of them with risk falling side effects, may lead to an increased risk due to the mixing of those drugs, possibly creating side effects that are not listed. The objective of this thesis is, among others, to identify some of those cases.

2.4.3 Risk factors

Since we are studying the impact on falls only in the elderly population we can associate most of the causes to that same factor, the age itself. Elderly people are most likely to fall because they have one or more of the following risk factors [Ton10]:

- being under the effect of certain medications, such as laxatives, diuretics, antidepressants or sedatives;
- suffering from conditions that affect balance, such as Parkinson's disease, arthritis, multiple sclerosis and stroke, or that cause sudden drops in blood pressure, like hypotension;
- having medical problems that result in the need to get up from bed at night or rushing, like insomnia and incontinence;
- using a walking aid, like a cane or walker;

- suffering from Alzheimer's disease or other forms of dementia

The age itself is a risk due to age related changes in the body. Problems like vision issues, balance, muscular problems, gait and psychological factors are all related to the age and will increase the risk of falling [Ton10]. Even during the fall event, a capable body will have a much better response to avoid or minimize an injury. Physical strength and reflexes are very important in the process and both are factors that may be compromised when we talk about a elder person.

2.5 Fall risk evaluation system

In order to evaluate a specific drug risk we have to assign a specific class to that drug. This drug evaluation has to be done according to a standardized pattern, so we can get information sources and has to be sufficiently specific to identify the subsets we want. After some investigation we found a fitting solution to this classification problem, the Anatomical Therapeutic Classification System (ATC).

2.5.1 Anatomical Therapeutic Classification System

The Anatomical Therapeutic Chemical (ATC) Classification System was created in 1976 by te WHO Collaborating center for drug Statistics Division [WHO13].

This pharmaceutical drug coding system seeks to divide drugs into specific groups according to specific aspects. Information like the organ/system they act on, therapeutic or chemical characteristics will be taken into account to evaluate the drug.

ATC can have more than one code per substance, since one specific substance can be used in several Anatomical groups:

> *Acetylsalicylic acid (aspirin), for example, has A01AD05 as a drug for local oral treatment, B01AC06 as a platelet inhibitor, and N02BA01 as an analgesic and antipyretic.*

Several different brand names can share the same ATC code, since the active substance is the factor that is taken into account when we attribute the ATC code to the Drug. The figure 2.1 illustrate the classification code for Acetylsalicylic acid divided specifically in the 5 ATC categories.

This classification system can be very useful since some of the drug types we have to identify are mid-levels like Antidepressants (N06A) which is a 3rd level sub class and we can identify all their subclasses as potentially risky.

2.6 Related projects

Currently, due to the increase on the number of people using smart phones, there are many studies concerning the fall detection that use the capabilities of these devices[KP09, CLC+11, DBY+10].

Table 2.1: Nonlinear Model Results

ATC CODE	Description#1
A	Alimentary tract and metabolism
A01	Stomacological preparations
A01A	Stomacological preparations
A01AD	Other agents for local oral treatment
A01AD05	Acetylsalicylic Acid

The reason for this is that the mobile phone is always with the subject and there is no need to get new equipment if the patient uses a smart phone. Nowadays smart phones have a vast array of tools capable of identifying possibles clues about the state of the patient. Using accelerometer, camera, compass or any other devices available, a smart phone can process the data and brings up a conclusion. Today there are mobile phones with more computing power than the average PC, which makes this process very fast and easy to implement.

The detection of falls is a part of several projects at FhP, but the main objective in this dissertation is not related to the fall detection. Instead we will analyse and reach conclusions about the medicine currently being used by the patient. We will evaluate the list of drugs currently used by the patient, place the drug in a certain class and from there assess the fall risk associated with the drug.

There are several studies regarding multi medication and our research revealed that some of these projects have similar objectives and methodologies. However, we did not find any particular initiative that is developing a system with the exact same objective, a polypharmacy study related with falls in the elderly population.

Instead we found several projects that use the mobile phone to evaluate ambulatory or health status on the patient based on the phone sensors (applications like "Smart fall Detection" or "Fall detection v1.0"). A specific study and concrete application was developed in Bolgana, Italy [TMC11]. We also found several applications evaluating or informing about drugs, but none of those applications have the specific objective to relate the medicine with falls.

2.7 Overview

With the literature review concluded, we can identified a clear problem to the elderly. The existing information regarding drug classes, side effects and drug-drug interactions allow us to create a automated system that can analyse the information and assess the fall risk to a specific situation. However, according to our research, such a system was not created by anyone with this specific objective of evaluating the risk inherent to the multi drug usage and falls. This system would diminish the possibility of human mistake when prescribing drugs to a patient, since it could be used as a tool that would check for possibly risky situations, and would save time to the healthcare professional.

Chapter 3

Resources and Tools

In this chapter we intend to present all the resources that were used along the development of this dissertation. First we will present the tool and technologies we worked on and then the resources and sources of information that were imperative in order to get the information we needed. All the tools and technologies will have brief description about their usage and relation to our work.

During the development of this fall risk evaluation system one of the most important challenges in this section is the decision of the tools, technologies and resources that should be used for the given purpose and scenario. During a brief period before starting the project we took some time in order to study the possible tools and technologies. All the data gathered at that time will be presented below.

3.1 Resources

3.1.1 Pharminx Ontology

Since our work is to evaluate the available drugs in the Portuguese market we needed a representation model of medical drugs and a source of information of the available medication. "Information Extraction From Medication Leaflets", a dissertation previously developed at FhP [Agu12] provided all that information and data structure. That dissertation aimed at automatically extracting information from medical leaflets and store all the retrieved data using a representation model. That model is capable of recording the data of the drugs and the relations between concepts. That representation model was named Pharminx and was used as the base our project. We updated and improved the existing model to use it in the evaluation process here. That model was already populated with the information extracted from the Portuguese medical therapeutic records [MH11]. The process of extracting that information was done using automated natural language interpretation, which may compromise the accuracy of the records. However, the results achieved during the test phase of the extraction process points to a high accuracy in the process of identification and

evaluation of the interactions with a success rate of approximately 97%. Those records contain many kinds of information about the pharmacological drugs currently being used in Portugal.

During the current project the Pharminx model was studied and updated in order to adapt it to the fall risk evaluation process. All records in the original population were also updated with all the data we had available in our sources. The figure 3.1 is a representation of the original data model. As we can see there are elements representing the possible data available in the Infarmed records[MH11]from where the information was extracted regarding drug and usage data. These records are available through the National Authority of Medicines and Health Products and the Portuguese Health Ministry. They contain many different kinds of information about all the pharmacological drugs currently being used in Portugal.

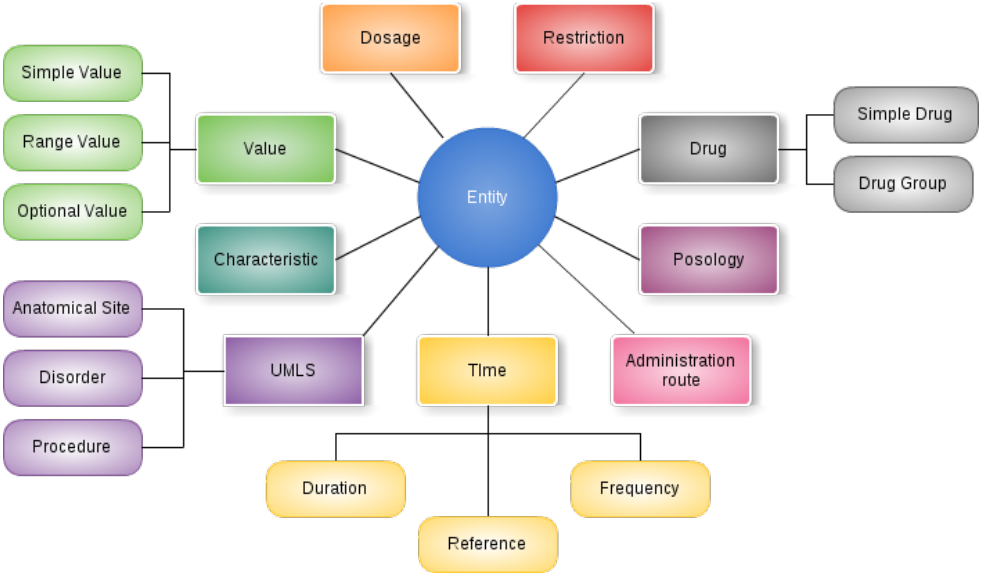


Figure 3.1: Original Pharminx Ontology Model

3.1.2 Unified Medical Language System

The Unified Medical Language System (UMLS) is a compilation of many medical related vocabularies with controlled sources. The UMLS can provide a mapped structure of those vocabularies allowing the user to check the relationship between the concepts and the translation of those concepts into dozens of languages. UMLS was developed by the United Sates National Library of Medicine (USNLM) in 1985 with the objective of encouraging the usage of all the diverse biometrical data available to electronic systems. The implementation of this system was a big step to the generalization of concepts, since it resolve, or at least minimizes the naming problem that exists due to the translation of concepts and medical drugs.

Currently UMLS stores over 1 million biomedical concepts and 5 million concept names spread over 100 incorporated controlled vocabularies and classification systems. This can be

accessed online using UMLS Terminology Services (UTS) or offline after installing the UMLS database [UA11]

Currently there are three major sources to UMLS database parts:

- **Metathesaurus** - The Metathesaurus forms the base of the UMLS and comprises over 1 milion biomedical concepts and 5 million concept names.
- **Semantic Network** - Used to represent the relationship between the concepts, among the semantic types that are available in Metathesauros.
- **SPECIALIST Lexicon,** - a database containing syntactic, morphological, and orthographic information for biomedical and common words in the English language.

For our work it is important that we can use and translate medical drugs available in Portugal. To do so it is crucial to check the availability of the Portuguese language in the UMLS Database. We used the latest version available, the 2012AA. According to the documentation this version uses 4 Portuguese sources listed below:

- ICPCPOR (ICPC Portuguese) - 1999AA
- MDRPOR (MedDRA Portuguese) - 2012AA
- MSHPOR (MeSH Portuguese) - 2012AA
- WHOPOR (WHOART Portuguese) - 1999AA

Using this UMLS version we were able to identify a total of 10,810,680 concepts each one of them with an individual Concept Unique Identifier (CUI) and 159,911 Portuguese translations available.

3.1.3 Anatomical Therapeutic Chemical Classification System

The Anatomical Therapeutic Chemical Classification System (ATC) was developed and established by the World Health Organization (WHO). In this classification system, all drugs are grouped in five distinct levels (anatomical main group, therapeutic subgroup, pharmacological subgroup, chemical subgroup and chemical substance). Possible modifications on the ATC database are only made by WHO specialists in their update releases and are based only in scientific articles and international pharmacologists [1].

> *"The drugs are divided into fourteen main groups (1st level), with pharmacological/therapeutic subgroups (2nd level). The 3rd and 4th levels are chemical/pharmacological/therapeutic subgroups and the 5th level is the chemical substance. The 2nd, 3rd and 4th levels are often used to identify pharmacological subgroups when that is considered more appropriate than therapeutic or chemical subgroups [WHO13].*[2]*"*

A	Alimentary tract and metabolism (1st level, anatomical main group)
A10	Drugs used in diabetes (2nd level, therapeutic subgroup)
A10B	Blood glucose lowering drugs, excl. insulins (3rd level, pharmacological subgroup)
A10BA	Biguanides (4th level, chemical subgroup)
A10BA02	metformin (5th level, chemical substance)

Figure 3.2: Representation of ATC subgroups [WHO13]

As we can see in the figure 3.2 we have an example of the ATC code for metformin. The five groups are rated separately and together they form a code describing the medical drug.

WHO has an available online query system to identify and classify the medical drugs available in their database. For this system we cannot rely on a system like that since there is no API nor the system is developed to this purpose, so we had to get our own ATC classification database. Since the ATC database was not free we had to search for alternatives. This search lead us to two ATC sources:

- Alphabetical Listing of ATC drugs and codes - a PDF file with 63 pages and a total of 3794 entries from WHO dated from 1999 and published by MSupply, a pharmaceutical company[Dro00].
- German Institute of Medical Documentation and Information - a PDF/XLS file with a total of 7602 entries from the Institute dated from January 2012 [WId12].

Both sources are of a considerable size, reliable and complete.

3.1.4 Side Effects Resource

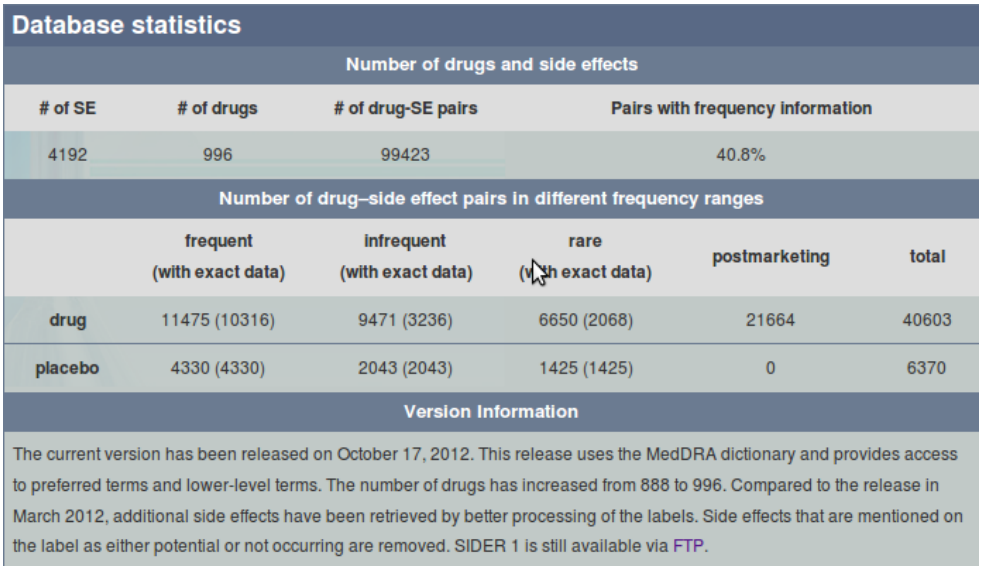
The SIDER Side Effects Resource represents an effort to aggregate dispersed public information on side effects. In 2008, when the database was created no such resource exist in machine-readable form. The creation of this resource was motivated by many requests for data that Monica Campillos used during the creation of a paper on the utilization of side effects for drug target prediction [CKG+08]

It contains information on the side effects of several drugs and was extracted from public documents and leaflets. The available information includes side effect frequency, drug and side effect classifications as well as links to further information, for example drug–target relations.

The SIDER Side Effect Resource Database will only be used to complement the information we already have available on side effects. It is an "extra" that can improve the analysis of fall risk

[1] http://www.whocc.no/atc_ddd_index/
[2] http://www.whocc.no/atc/structure_and_principles

related to a specific drug and allow us to get more results during the drug evaluation phase. We used the latest version available dated March 2012.



Database statistics

Number of drugs and side effects

# of SE	# of drugs	# of drug-SE pairs	Pairs with frequency information
4192	996	99423	40.8%

Number of drug–side effect pairs in different frequency ranges

	frequent (with exact data)	infrequent (with exact data)	rare (with exact data)	postmarketing	total
drug	11475 (10316)	9471 (3236)	6650 (2068)	21664	40603
placebo	4330 (4330)	2043 (2043)	1425 (1425)	0	6370

Version Information

The current version has been released on October 17, 2012. This release uses the MedDRA dictionary and provides access to preferred terms and lower-level terms. The number of drugs has increased from 888 to 996. Compared to the release in March 2012, additional side effects have been retrieved by better processing of the labels. Side effects that are mentioned on the label as either potential or not occurring are removed. SIDER 1 is still available via FTP.

Figure 3.3: SIDER database statistics

As we can see in the figure 3.3 this database has a total of 4,192 side effects available linked to a total of 996 medical substances and respective trade names, each one of them connected to the UMLS via a Concept Unique Identifier (CUI). This will allow us for an easy connection to the remaining data structures in the project since we already have the CUI on all the identified concepts via UMLS.

3.1.5 Drug-drug interaction information source

To evaluate the interaction between drugs we must find some reliable source of information that is able to specify the reaction between two or more distinct drugs. During our research we found some options that could be used as source. Most of the sources we found were either a private source or a product for sale. Those circumstances form an obstacle to the development of our drug interactions system. In some cases we contacted the company/individual responsible for the product to see if it was possible to get partial access to the databases. However we did not manage to gain access to the resources. Options will be discussed individually for each of the sources we found and the reason why we use them or not.

Sfinx drug Database

Swedish, Finnish, INteraction X-referencing (SFINX) is a drug-drug interaction database that belongs to Medbase [3] LTD in Turku, Finland, the Karolinska Institute Department of Clinical Pharmacology in Stockholm, Sweden, and the Division of Drug Management and Informatics at

[3] http://www.medbase.fi

Stockholm County Council. This database contains concise and validated information concerning consequences and recommendations on the combination of drugs for more than 12,000 interaction pairs.

This database also contains the substance names, Chemical Abstract Service (CAS) number, ATC code and some other informations. SFINX is distributed as an SQL database which is updated four times a year and delivered as XML-format export file.

This solution seems optimal to our system, but after exchanging e-mails with the responsible at Medbase we were informed that there are no Distributors available in Portugal and even if there was we had no guarantee that the database would be available freely.

Lexi-Comp Drug Interaction database

Lexi-Comp Inc.[4] is a company that develops clinical information solutions. It also provides other tools to software development, improve patient safety and elevate the quality of care patients receive. Lexi-Comp provides for customers a Data list of drug-drug, drug-allergy and drug-disease interactions updated regularly. This data is integrated into an Electronic Medical Record (EMR) and can be accessed through a Web API Solution, a Web Service or XML Datasets. However it has the same problems as SFINX since it is property of a private entity and can not be accessed freely. In this case we also tried to contact with Lexi-Comp but without results.

Drug Interactions Checker

Drug Interactions Checker [5] (DIC) is an online service provided by Cerner Mulum, INC. It is a service with the objective of easily and rapidly provide interactions information about a vast array of medical drugs to healthcare professionals and usual customers/patients. The problem with this source is that there is no available API or access form to the data in the format we want it. Besides, the information is only available online and cannot be stored locally, which is a problem for the type of system we want to develop. This service was created with the objective of simplifying the process of checking for drug interactions for the public with a fairly simple and free access.

Drug DDI Corpus

The DrugDDI Corpus is part of a project developed by Isabel Segura-Bedmar and Paloma martínez with the association of the University Carlos III of Madrid [SBMdPS10]. "*The DrugDDI corpus is part of a larger study about automatic Drug-Drug Interaction Extraction. The corpus provides data for the development and automatic evaluation of systems that annotate and extract drug-drug interactions*" [6].

This project uses textual information input regarding the drug combinations obtained from the DrugBank database [7]. This database is a rich resource that combines chemical and pharmaceutical

[4] http://www.lexi.com/
[5] http://www.drugs.com/drug_interactions.html
[6] http://labda.inf.uc3m.es/DrugDDI/DrugDDI.html
[7] http://www.drugbank.ca/

information regarding approximately 4,900 substances.

The corpus is based on a set of 579 texts describing DDI's(Drug-Drug Interactions) from DrugBank database. The texts were manually annotated with the interactions with the assistance of an expert pharmacist, which ensures for the credibility and veracity of this source. A total of 5,806 sentences were evaluated, containing a total of 3,160 Drug-Drug Interactions.

The result is available to public as XML documents organized by drug names/chemical substances regarding the interpretation of each individual DrugBank document in the corpus.

Here we have an example of a sentence correspondence between the text and XML data from the corpus:

> *In vitro studies also indicate that celecoxib, although not a substrate, is an inhibitor of cytochrome P450 2D6*

Resulted in the XML portion shown in figure 5.4

```
−<sentence id="DrugDDI.d53.s4" origId="s4" text="In vitro studies also indicate that celecoxib, although
 not a substrate, is an inhibitor of cytochrome P450 2D6.">
   <entity id="DrugDDI.d53.s4.e0" origId="s4.p45" charOffset="36-45" type="drug" text="celecoxib"/>
   <entity id="DrugDDI.d53.s4.e1" origId="s4.p50" charOffset="92-111" type="drug" text="cytochrome
   P450 2D6"/>
   <pair id="DrugDDI.d53.s4.p0" e1="DrugDDI.d53.s4.e0" e2="DrugDDI.d53.s4.e1" interaction="true"/>
 </sentence>
```

Figure 3.4: XML representation example from Drug DDI corpus

According to the available information this data source is reliable and is the bigger annotated corpus for DDI that currently exists. One of the disadvantages we have with this set is that the other possible sources mentioned about above have more entries and drug interactions and most of them include a SQL database that is not available in this option. However, since this source is completely free and available we will use it in our project. The first step in the usage of DrugDDI will be the interpretation and parsing of the xml files to a easier access format. In this case a SQLite database.

3.2 Tools and Technologies

3.2.1 Web Ontology Language

The Web Ontology Language (OWL)[8] is designed to be used in applications that need to process and have some sense of understanding the information content available. In this description language we have specific relationships and dependencies between the data parts of the model. OWL facilitates the machine interoperability and uses XML, RDF and RDF-S vocabulary. This kind of Language is written as XML document and uses the same N-triples and Turtle formats as RDF. We can divide OWL into 3 sub languages listed below:

[8] http://www.w3.org/TR/owl-features/

- **OWL Lite** - supports those users primarily needing a classification hierarchy and simple constraints.

- **OWL DL** - supports those users who want the maximum expressiveness while retaining computational completeness and decidability. OWL DL includes all OWL language constructs, but they can be used only under certain restrictions.

- **OWL Full** - it is the most advanced and complete OWL and is meant for users who want maximum expressiveness and the syntactic freedom of RDF with no computational guarantees.

3.2.2 Protégé

Protégé[9] was developed by Stanford Center for Biomedical Informatics Research at the Stanford University School of Medicine. It is a free, open-source platform that provides a tool to construct domain models and knowledge-based applications with ontologies. Here we will use it to edit the ontology structure. This means adding or removing properties between elements of the ontology. It implements a rich set of knowledge-modelling structures and actions that support the creation, visualization, manipulation or edition of ontologies of various representation formats.

To develop this thesis we used one of the last available and stable versions of Protégé, the 4,1 since this was the version used in the development of the original Ontology.

3.2.3 Jena

Apache Jena[10] is a Java framework developed for building Semantic Web applications and in this case we will use it to edit the OWL population. It provides a vast collection of tools and Java libraries to help the developer to work with semantic web and linked-data applications like the one we will develop here. In this specific case Jena will be used to manipulate and access the information available in OWL.

Jena is developed so it can cope with this description language in an organized and effortless way for the developer.

Jena is also capable of executing a rule based inference over the RDF or OWL sources. It allows the user to store a large number of triplets/data and be efficient at the same time. One of the main advantages of using Jena is that it is compatible with SPARQL, which allow us to do fast queries on the model without having to do a manual search that would take much more processing time.

[9] http://protege.stanford.edu/
[10] http://jena.apache.org/

3.2.4 MySQL

MySQL[11] is largely known as the world's most used open source relational database management system. Here we will use it to store all the UMLS information, specifically, the concepts with the respective Concept Unique Identifiers, translations and relationships. This is one of the distributions forms of UMLS, and our choice since it is effective, free, and allows for fast and extensive queries.

3.2.5 SQLite

SQLite [12] has a fairly similar structure when compared to MySQL. It uses the same query syntax and table/schema organization. However, there are a few specific aspects that make SQLite the perfect database for small and local databases. First of all SQLite databases will be stores in a single file and only one user at the time can have access. The performance of these databases is similar to MySQL's when we are using small databases as the ones we will use in this project.

We will use SQLITE to store the information extracted from PDF, TSC and XLS files containing ATC, drug-drug Interactions, UMLS subsets divided by languages and side-effects data.

[11] http://www.mysql.com/
[12] http://www.sqlite.org/

Chapter 4

Knowledge Representation Model

As mentioned in the introduction of this thesis, this project was developed under the auspices of Fraunhofer Portugal AICOS. The project objective is the study of the relation between elder persons, multi-medication and falls using the technologies available to the society today, to help in the study and prevention of the falls in the elderly population.

Our part in the development of this system is the design and development of a tool that is capable of analysing the medication currently being taken by a specific patient and evaluate the possible fall risks associated with that medication. This evaluation can be made by analysing either the drug type/family, side effects or a specific drug combination.

In this chapter we intend to present a more technological overview of the knowledge representation model we are going to use and upgrade in the development of this project. Note that the implementation of the project was done using as base a previous ontology developed here at FhP by Bruno Aguiar [Agu12] during his dissertation. We studied the base ontology used on that project in order to use it and update it with fall related data.

4.1 Original Model

The knowledge representation model will be used to store the knowledge we have on the variables of our system. The original data model has a lot of data that we do not need, but there is no need for us to delete the relations or the data. We can simply add the parameters we need and edit the population updating the records with new information we will use to assess the fall risk. This will allow for a more complete and extended representation model that can be used with other objectives related to the extra information available. The original model, represented by the Figure 4.1, contains Posology, UMLS records, Time variables, Administration routes, Drugs and Drug groups, Dosages, characteristics and Values. The only pre-existing fields that we will have to use here are the Drugs and UMLS CUI records. The CUI is an important information because

it will allow us to connect the concept in the ontology with all the information it has available in the UMLS records.

The knowledge representation model will not represent the information in a usual way as it is represented, for example, in a simple database. It will represent the relations and purposes of each attribute and their dependencies.

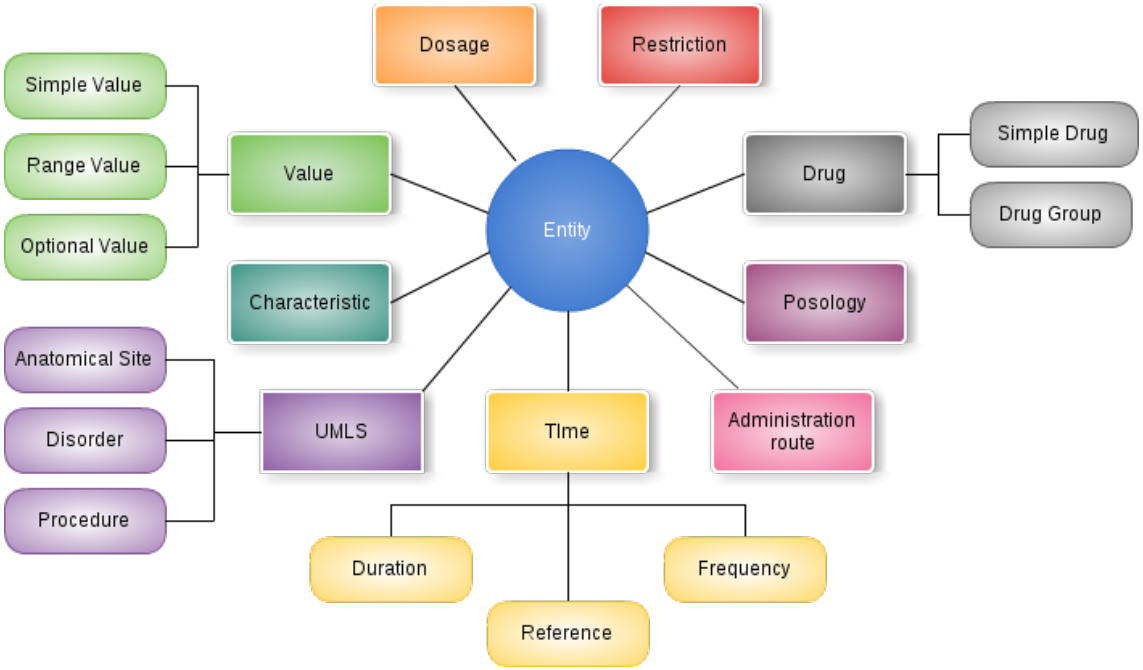


Figure 4.1: Schematic knowledge representation model of the original ontology

Knowledge modelling is a process where the information becomes interpretable by a computer using a model of knowledge or standard specification about a certain kind of process limited to a predetermined field, in this scenario, the drugs and related information. This knowledge representation language or data structure is the only way for the information to be computer interpretable. We can quote Sowa who says:

> *"Knowledge representation is the application of logic and ontology to the task of constructing computable models for some domain"* [Sow00]

In the original model, represented in Protégé illustrated by the Figure 4.2, we can see that we already have several classes that may be used in our process.

- **DrugDescription** - This class will be one of the main pillars in the development of our project. We will use the existing structure, mainly in the subclass Drug and add some new structures that will be connected to this class.
- **UmlsEntity** - Here we can connect the concepts with the entity available in the UMLS database. The subclass we will be using in this case will be Disorder. It will be connected

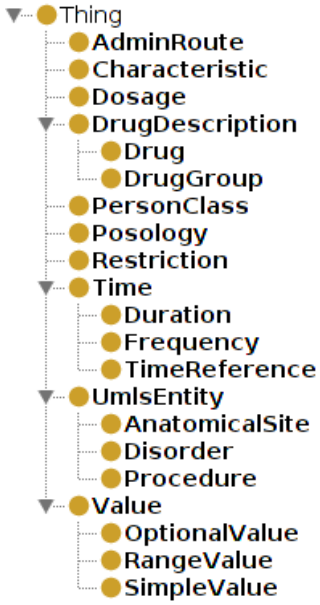


Figure 4.2: Knowledge representation model of the original ontology using Protégé

to the Drug as a description to the side effect. All disorders that will be listed as side effects already have CUI's that will be useful in the evaluation process.

These classes will be very useful in the fall risk evaluation, however, to make a more precise assessment of the risk we will add several new data structures in the model.

4.2 Updated Model

The Figure 4.3 represents the specific fields of the ontology that we worked on. All the necessary changes and updates to the structure of the data model are presented on Figure 4.3 and are then explained. Note that the disorder class was used when we updated or added side effects.

The update we applied to the model was very specific to some parts that we verified before to be related with the fall possibility. First of all, and as support field to our work and any other work that needs to gather information we decided to add the CUI field to the DrugDescription class. This will help us not only to get the translations to the concepts we need, but will also allow for us to get several types of information we may need, such as related concepts, brand names or even components of a drug.

One of the big changes on this specific part of the ontology is the addition of various Data Properties regarding the ATC code. Therapeutic SubGroup, Anatomical main Group, Pharmacological Sub Group, Chemical Sub Group and Chemical Substance were the 5 fields added regarding ATC. These new structures and information will be useful in the identification and interpretation of the drug group or family, and this information can be directly associated with a relative fall risk. The reason why we decided to add these parts of the ATC separately is because it may be useful for us in some cases to get only the sub families of a drug.

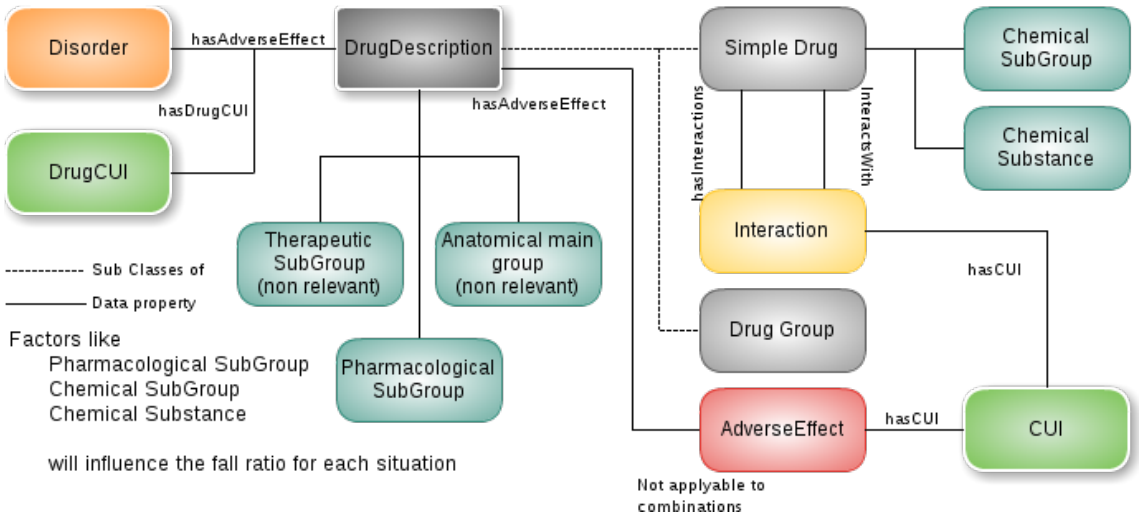


Figure 4.3: Ontology modifications/updates

One other part were we did not change the structure, but updated the data was the Disorder class. We used this class when adding side effects based on the SIDER database. It is connected to the Drug class by the "hasAdverseEffect" object property. We also added for each new Disorder entry the respective CUI. This enables us the translation from English available on the original source to Portuguese, the ontology language. It also allows us to check if the Disorder already exists in the original ontology. If the disorder already exists we only add a new possible name to it.

A final but even more important addition to the ontology is the "hasInteraction" object property. This property will be used as a representation in the drug-drug interactions. It is a property that is responsible for handling the interactions of a specific drug. If the interacting drug is already present in the ontology this class will be directly connected to it. Otherwise and since we do not want to add new drugs to the ontology population we will just add the CUI regarding the interacting drug. That will allow us to have a complete record of the interactions without adding new drug entries.

Chapter 5

Fall Risk Evaluation Process

The evaluation of a drug combination case has several variables that will influence the likelihood of a fall occurrence. We decided to divide the fall risk evaluation process in two specific parts. First the individual evaluation of a drug, regarding the identification of the drug properties, specifically family/group, related side effects and respective risk. After this process we will check for possible interactions between the prescribed medication to the patient.

5.1 Single Drug Evaluation Process

This part represents a very important step in the process of risk assessment. Here we will analyse the entry drug names and identify them with a specific data/concept entry. If the respective data entry for the specific drug is not available in our model we will access the entry by consulting our available sources, ATC, side-effects and UMLS, in order to construct a temporary model that will allow us to make a evaluation even if the ontology population data is insufficient. The ontology will only be updated with this new data if the entry already exists but new information is available. We choose not to add new concepts once the original source of the data model population only contains the Infarmed records of drugs available in Portugal and we do not know if we are adding a drug that respect that condition.

Figure 5.1 represents a scheme of the simplified Single Drug Evaluation Process (SDEP) outside the ontology model. This process will use most of the information sources available to us in the project.

The entry name can be given as a Chemical substance or a specific trade name of a pharmaceutical drug. The SDEP must be capable of conducting an evaluation process in both cases. Since the ATC, side-effects and UMLSSubset databases only use chemical substance as entries we must identify if the entry represents a trade name and retrieve the specific related active substance. To check the type of entry we received we will use the UMLS database. We opted for the direct usage of the UMLS dataset since the amount of information that would be necessary for us to make an

intermediate database would be too large to make a SQLite database to work effectively, and we could lose some entries in the process.

After checking for trade name exception we can finally associate the entry with a specific chemical substance Concept Unique Identifier (CUI). This code will allow us to translate the chemical substance to all the languages available in the UMLS database. The translation is important since the various information sources we have are from different countries, which means they have different languages.

The next step consists in the retrieval of the ATC code for the chemical substance. Having the language translations for a substance will allow us to use the ATC database search available in English and German.

An identical process as the one verified in the ATC code retrieval will be used to get the possible side effects on the substance usage. This step is optional, since most of the drugs in the original ontology populated using Portuguese Infarmed records already have some listed side effects. We use SIDER database to allow us for a more complete side effects listing.

The final step in the SDEP is the evaluation of all the data previously gathered in this process to get a conclusion regarding the fall risk related to the entry drug. This process will be explained step by step ahead in this section.

5.1.1 Entry Evaluation Process

The system must be able to face any possible entry and give an adequate response. Ideally, the entry in this case must be either a chemical substance or a specific trade name. To make this kind of analysis we must rely on the UMLS database. It is possible to verify this type of data since a large part of the concepts listed in UMLS have the "type" information. Unfortunately, the concept and the type are listed in two different tables. In a database like UMLS where the main table regarding the concepts contains roughly 10,800,000 entries joining tables would be a problem to the query performance. To do this in an effective way we have done a composed query. This means that we will do multiple queries instead of one query that would have required a join in the tables. Doing multiple queries may seem like a process that would consume a lot more time, but since none of these new queries will require a join the execution time will be much faster when compared to a single query that would do all the work.

The tables 5.1 and 5.2 represent the relevant fields in the tables MRREL and MRCONSO from the UMLS database.

Column	Description
CUI1	First Concept Unique Identifier for chemical substance
CUI2	Secondary Concept Unique identifier for Trade Name
RELA	Relation type between the concepts. In this case "has_tradename"

Table 5.1: MRREL relevant content

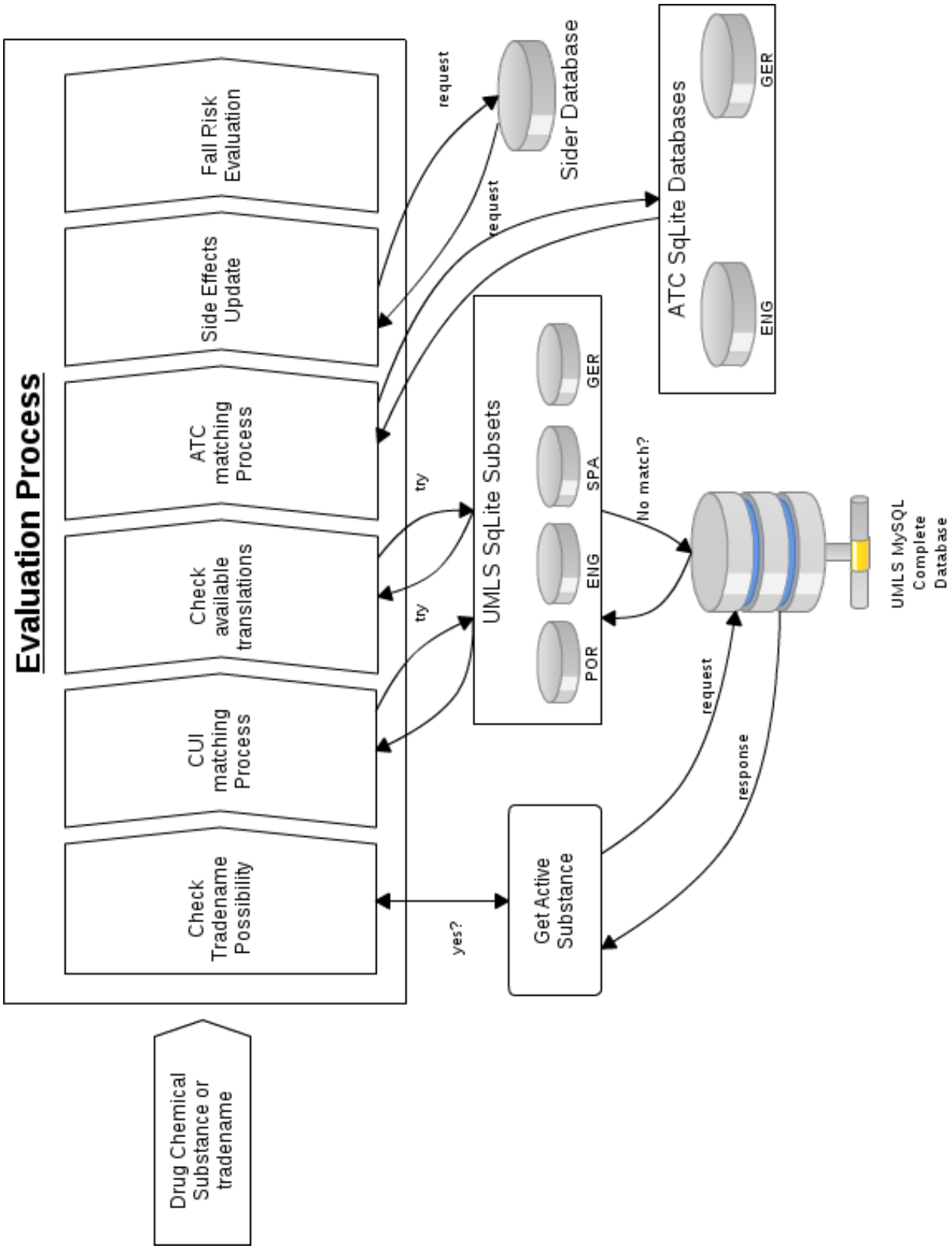


Figure 5.1: Project Layout

Column	Description
CUI	Concept Unique Identifier
STR	String or concept name
LAT	The Language of the concept (POR, GER, ENG, SPA, etc...

Table 5.2: MRCONSO relevant content

With this information available we can execute simple queries to know the type of the entry. The queries shown in Listing 5.1 will access the data available in the tables without the need to execute a join. In this example we use the trade name "Prosom", used for the chemical substance "Estalozam".

```
1
2  Select CUI
3  From MRCONSO
4  Where STR = "Prosom";
5  +----------+
6  | CUI      |
7  +----------+
8  | C0722865 |
9  +----------+
10
11 Select CUI2, CUI1 RELA
12 From MRREL
13 Where CUI2 = "C0722865" AND RELA = "has_tradename";
14 +----------+----------+---------------+
15 | CUI1     | CUI2     | RELA          |
16 +----------+----------+---------------+
17 | C0722865 | C0014892 | has_tradename |
18 | C0722865 | C0014892 | has_tradename |
19 +----------+----------+---------------+
20
21 Select STR
22 From MRCONSO
23 Where CUI="C0014892";
24 +---------------------------------+
25 | STR                             |
26 +---------------------------------+
27 | estazolam                       |
28 | Estazolam                       |
29 | ...                             |
30 +---------------------------------+
```

Listing 5.1: Concept query

After understanding the operation to get the trade name relations this query is fairly simple. First we get the CUI , if it exists, for the entry name. After that, we use that same CUI to know

if there is any chemical substance linked to the given trade name. As we can see in the example above, the CUI2 represents the chemical substance linked to the entry trade name.

This CUI2 represents the connection we want to the entry. At this point we can get all the information available in the UMLS database regarding this substance/concept. That information will be very useful to evaluate the drug entry since it will allow us to translate the drug name into various languages, a process that can be used to connect the entry with ATC, side-effects or drug combinations that may be available in other languages.

5.1.2 The UMLS SQLite Subsets

As we said before, the UMLS database is the existing database regarding medical concepts and relations. It compiles all the information available in various reliable sources into one single and organized database. This means that we will have on the UMLS database a lot of content that we will not be using and that content may cause a decline in the query performance. Facing this problem we decided to create subsets of the existing database separated by languages. Another reason for the creation of this subsets is the disk space usage for the UMLS database. It is possible for this system to be implemented as a web service. The disk usage would be a problem then, mainly if we talk about regular backup management and creation. All these factors lead to the decision of creating subsets of the UMLS database that can be used in some cases to get faster results and smaller disk usage. We should mention the fact that we can use the system in both ways, either by using directly and only the UMLS MySQL database or by using only the SQLite language subsets.

It would be also possible to create a UMLS database using only some types of concepts or languages. During the development process we also tried this approach that is an option provided by the UMLS installation package. However, after limiting the content types and languages we still identified some concepts missing when compared to the entire set. This may be due to some error in our configuration or some problem in the database construction process. Since we want to have the maximum matching possible we decided to use the complete UMLS dataset and enhance the query performance to the best possible point.

Analysis on the matching for both scenarios, UMLS full database and UMLS subsets, will be presented in the Result evaluation section. The Table 5.3 represents the structure we used in the SQLite language subsets and an example for each field regarding the substance *Mesna*.

Column	Example
CUI	C0000294
NAME	Mesna
LAT	POR

Table 5.3: UMLS SQLite subset structure and example

A total of 4 subsets were created with fragments of the MRCONSO table that will allow us for a faster query on concept names and CUI, according to the language. The original UMLS MySQL

database version we worked on here uses approximately 25GB of disk space. The 4 subsets we have created (Portuguese, Spanish, German and English) all together will use 427 MB. These subsets will be used in the translation of concepts needed in order to connect the other available sources of information. The UMLS query for translations will only be used if it is impossible to find the concept in the subsets.

Database	Size (MB)
UMLSPTSubset	1
UMLSENGSubset	362
UMLSPASubset	64
UMLSGERSubset	1.2

Table 5.4: UMLS SQLite Subsets space usage

As we can see in the Table 5.4 there is a lot more information available in the UMLS database in English, mainly when compared to the Portuguese language. This is due to the largely higher number of English medical sources.

5.1.3 The Anatomical Therapeutic Classificatisuperioron System Matching

The ATC code is one of the main evaluators we use in the risk assessment of a drug. This code will be the bridge between a simple concept name or identifier to a code that we can interpret to know the kind, family or type of drug we are dealing with.

As we mentioned before in the Resources and Tools section, we used two different sources for the ATC code entries.

- Alphabetical Listing of ATC drugs and codes - a PDF file with 63 pages and a total of 3794 entries from WHO dated from 1999 and published by MSupply, a pharmaceutical company.
- German Institute of Medical Documentation and Information - a PDF/XLS file with a total of 7602 entries from the Institute dated from January 2012.

The source form of the documents does not allow us for an easy machine readable solution, so, in both cases, we had to parse the information into a standard form. Since we already had an SQLite interface ready to use we decided to convert both files into two SQLite Databases with the structure shown in Table 5.5.

Name	ANATOMICAL	THERAPEUTIC	PHARMACOLOGICAL	CHEMICAL GROUP	CHEMICAL SUBSTANCE
Abciximab	B	01	A	C	13

Table 5.5: ATC SQLite database schema

As we can see, the structure for the ATC data will divide the code into 5 parts regarding anatomical, therapeutic, pharmacological, chemical group and chemical substances. Besides the 5

ATC parts we will have the name in the source language that will serve as connector to the entry. The matching process will try a direct match between the translated name and the database entry. If this does not work it will try another approach for matching that we will explain later on this section.

At this point and assuming we found a match we already have the CUI for the chemical substance and consequently we have several translations that allow us to make the connection with the requested ATC.

5.1.4 Side Effects Database

As we mentioned before the side effects have a unique and important influence in the probability of a fall occurrence. The original ontology population already had a fairly complete list of possible side effects related to certain drugs extracted from the Infarmed records. These side effects are listed as Disorders, a subclass of UmlsEntitys. However, some drugs had no listed side effects and in order to manage a more complete evaluation of the fall risk we decided to get a source that could complete the data for the substance's side effects. That is where we used the SIDER Database mentioned in Chapter 2.

Once again, this is an option, since we can evaluate the drug using only the available side effects that are already in the ontology population. This database is a way for us to manage an evaluation on the side effects if the drug is not found in the ontology population or if the entry in the ontology does not list any side effects. It can also be used to update the existing side effects on the ontology population.

The SIDER database was available as 3 tab separated values (tsv) files that we once again parsed into a SQLite database. This database contains generic and brand names of the medical drugs, URL, for the sources used, UMLS CUI's for the side effects and side effect names, all in English. It also contains a final .tsv file that has information on the frequency of the side effects. However, that information is very incomplete for some drugs, so we did not use this part of the database.

The structure adopted for the SideEffects database is listed in Table 5.6. There we can also see an exaple of one entry for the substance *levobunolol*.

Column	Example
DrugName Database schem	levobunolol
SideEffects	cerebrovascular accident+rash+...
CUIs_SideEffects	C0038454+C0015230+C0033377+...

Table 5.6: SideEffects SQLite fields and example

Another aspect we must reinforce is that this side effect database only contains side effects of specific substances or trade names. In the ontology, we often find drug entries with substance mixing. This means that we cannot just apply the information in this database to those cases. The mixing of two substances will not necessarily result in the mixing of the side effects from both

substances. So, in case of a multi substance entry, the side effects query will not be made and we will not add data to the ontology model since we have no way to assure the side effect's outcome in that scenario.

5.1.5 Matching process and Levenshtein distance

During all this process several string matching operations had to be done regarding drug trade names and mainly substances. The process used to do this match will influence the accuracy, matching ratio and efficiency of the process. We can say that per substance we can have normally 6/7 matching processes in the drug evaluation and information gathering process. Here we explain the exact process of matching for these names. The most common parts are listed in Table 5.7 .

Source	Matching	Purpose
System Entry	Ontology	to find ontology entity
System Entry	UMLS/UMLS subsets	to get CUI
System Entry	ATC Databases	get ATC in case CUI not found
Ontology	UMLS/UMLS subsets	to get translations/CUI
UMLS/UMLS subsets	ATC	get ATC
UMLS/UMLS subsets	SIDER	get SideEffects
UMLS/UMLS subsets	DrugDDI	Get drug interactions

Table 5.7: Matching process phases

We tried to use the same process every time we needed to link string entities. The process is designed to get the maximum matching possible and be as time efficient as possible. To get a precise match we first try to apply a translation on the entity. This step is impossible when we are dealing with a name that has not yet been linked to the element itself or its CUI, so we can not translate. In the other phases we get the array of translations available and try a direct matching. Figure 5.2 represents the work flow on the first step of the matching process.

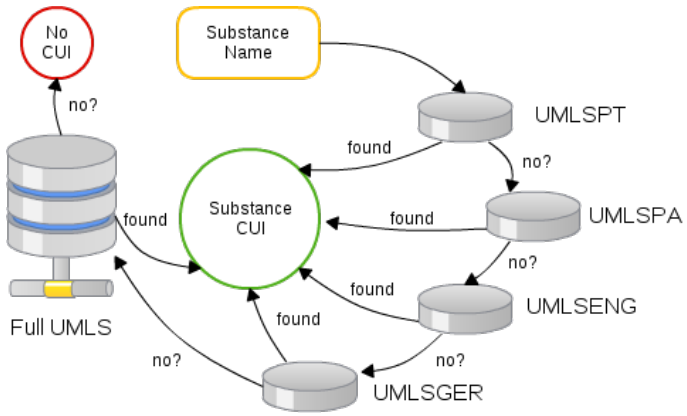


Figure 5.2: CUI search with direct matching

At first we use the UMLSPT subset since it is the smaller, faster, and the most likely to correctly translate in the scenario we have. If we do not find a match in the UMLSPT subset then we try the Spanish, English and German subsets and only as last resort we use the full UMLS dataset.

If we still have no match we will try to get a match using the Levenshtein distance. This resort is a bit more time consuming when compared to the attempts we have described before since it requires more computational power.

The Levenshtein distance is a string metric measuring system to check the existing difference between two character sequences. It measures the minimum number of single character edits required to change one word into another. So basically, if the two words are the same the Levenshtein distance will be 0. Figure 5.3 represents the mathematical formula to calculate the distance for two words.

However, this method, as a word difference measurement, may have an issues. The result of the method will be only one number that does not take into consideration the original size of the source words. To solve this problem we used a fairly simple approach: a normalized edit distance (NED) [Hee04, KD04]. The calculation of NED 5.1 takes into account the size of both words and will return a percentage value. This is a simple method where the computational power required is little compared to the alternatives. It also is one of the best choices we have on this scenario since the translations or variation of the names will typically differ in 1 or 2 characters, which comprises an already close match.

$$\mathrm{lev}_{a,b}(i,j) = \begin{cases} \max(i,j) & , \min(i,j) = 0 \\ \min \begin{cases} \mathrm{lev}_{a,b}(i-1,j)+1 \\ \mathrm{lev}_{a,b}(i,j-1)+1 \\ \mathrm{lev}_{a,b}(i-1,j-1)+[a_i \neq b_j] \end{cases} & , \text{ else} \end{cases}$$

Figure 5.3: Levenshtein Distance

Mathematically, the Levenshtein Distance between two strings *a,b* is given by lev $_{a,b}$(|a|,|b|). To simplify we show the following example using the words *kitten* and *sitting*.

```
1 ----- kitten -> sitten  (substitution of "s" for "k")
2 ----- sitten -> sittin  (substitution of "i" for "e")
3 ----- sittin -> sitting (insertion of "g")
```

$$NED = \frac{BiggerWordSize * lev(Word1, Word2)}{BiggerWordSize} \tag{5.1}$$

If none of those methods succeed then we can not connect the entry to the UMLS concept. Even in this case scenario we can proceed to try the ATC and side effect matching, but it will most likely achieve no results.

Once we have the CUI we can do the rest of the name matching sequence with the original name and available translations, prioritizing the translations in the same language as the source of information. This will increase the effectiveness and performance in the matching process.

5.1.6 Drug Evaluation Process

Once we have gathered all the possible information available for each drug we will evaluate it regarding the risk of falling. We can do this evaluation for a single drug in two separated parts. The first part is related to the ATC interpretation. Since the ATC code allows us to verify the drug types, family and group we will check if that drug is a part of any risk groups and evaluate the inherent risk.

In order to do this we created a subset of ATC codes or parts that might be representing the risk drugs. The other part regards the side effects evaluation. In this part we will check the side effects that are connected to falls and reach an evaluation based on the number of fall risk side effects.

5.1.6.1 ATC Evaluation

In this section we will attribute the ATC code to the drug types listed as risky and explained in Chapter 2. Each entry to this set will have the related ATC codes and a reference to the article or documents that study and explain the connection to the possibility of falls. The Table 5.8 represents a list of the ATC codes we will take into consideration.

Family/Type/Group	ATC code
Bezodiazepines	N03AE, N05BA, N05CD, N05CF
Antidepressants	N06A, N06AX, N06CA
Selective serotonin ruptake inhibitors	N06AB
Tricyclic Antidepressants	N06AA
Antipsychotics	N05A , N05C, N05B
Antiepileptics	N03, N03A, N03AX
Central Nervous System (CNS)	V09A
Antihypertensives	C02
Beta-Blockers	C07 , C07A , C07B, C07C, C07D, C07E, C07F
Muscle relaxants	M03
Opioids	N02A, N01AH

Table 5.8: List of fall related drug families/types/groups

The ATC code from the drug currently being evaluated will be matched or partially matched to the closest subset of the listed ATC codes. If a match is found we know that the drug is listed as a fall related drug. The level of risk will depend on the match we have found. For example, we know that Benzodiazepines represent on of the riskier drugs and other drugs like Opioids have some inconclusive studies.

Each entry in this ATC list have a small description associated that will be displayed if the drug is a match. That small description will explain the reason why the drug is potentially risky and some additional information.

For example consider the active substance *mesoridazin*. The ATC code retrieved was N05AC03. In the ATC risk subset we can find a partial match with N05A. This will tell us that mesoridazin is an Antipsychotic drug and that there is a fall risk associated with it. The message associated with mesoridazin will be:

> *N05A -> Antipsychotic drug family This group comprises drugs with antipsychotic actions (i.e. neuroleptics)*

5.1.6.2 Side Effects Evaluation

As mentioned before, we can assess the side effects related to a specific chemical substance. However it is not relevant to our purpose to list all the side effects that may be caused by the usage of that substance. In order to study the relevant side effects we conducted a test.

First we tried to find any official and complete document reporting the most common fall related side effects, however it was not possible. Nevertheless, we found a document provided by the Mansfield District Hospital that listed some side effects that we took into consideration[Hos09].

As an alternative we used a list of fall related drugs. The objective of the test is to evaluate the recurrence of side effects connected to those drugs. The list of fall related drugs was provided by UNC-Chapel Hill news release [Etc08]. In this list we have a total of 93 high risk drugs from which one of them was not found and eight were not detected as risky by our system. Most of the failed detections were due to the Risk ATC not being listed as risky. In the next step we gathered all the listed side effects on the identified drugs.

A total of 15 drugs where not listed in the side effects database, so we could not associate their side effects to this study. We took the side effects with 35+ re-occurrences and studied them in order to assess their relationship with falls. We will conduct a more extensive test on the matching in the Chapter 6 of this document. The reoccurring side effects and the number of recurrences are listed in Table 5.9. Note that during the process of side effect identification a total of 78 different side effects linked to the drugs were found.

Note that the side effects listed in Table 5.9 are only the ones with 35+ occurrences and that according to their description may be linked to falls. Side effects like vomiting, nausea, rash and some other were not listed since they have no direct connection to falls. We used medicinenet.com [1] to evaluate the side effects type, description and relevance to this case scenario. Medicinenet is an online healthcare media publishing company. It provides easy-to-read medical information for consumers. It is directed by a experienced team of qualified professionals in various fields of medicine.

[1] http://www.medicinenet.com

Side-effect	Number of occurrences
headache	65
hypotension	64
confusion, ataxia,dizziness	62
edema	59
anorexia,syncope	58
tremor	57
hallucinations	56
blurred vision	50
somnolence, euphoria, insomnia	47
weakness, vertigo	46
drowsiness, dyspnea	45
abdominal pain	44
bradycardia, hypertension, hyperkinesia	42
incontinence	41
dystonia, incoordiantion, dyskinesia	40
diplopia, arthralgia	39
asthenia, postural hypotension, photosensitivity, psychosis	38
cramps	37
back pain, urinary frequency	36
peripheral edema	35

Table 5.9: Most common side effects for falls

We gathered a total of 102 side effects that have a clear relation to the fall risk of a medical drug. The side effects listed were taken from the study we have conducted and by literature research [EBM02, RL93, GHG+10, HLL07].

Using our side effects source database we can have access to the side effect's CUI. This will allow us to do a connection with the side effects in the original ontology population and will avoid the overlap of existing side effects.

Using this method we will list only the relevant side effects for each case.

5.2 Drug-Drug Interaction Evaluation Process

To do this kind of evaluation we must do the same drug entry identification process as the one we do for single drug evaluation. The translations available will increase the possibility of finding a match and consequently the accuracy of the result. As mentioned before we used the DrugDDI Corpus as a source for the drug-drug interactions. We had to convert the XML source files into a SQLite database. Figure 5.4 represents an example of the original format of the information.

We only took into account the drugs where the Interaction tag was true. From there we only had to link the entity ID to the drug name listed in the XML as well. The Objective of the database is to contain the names of both drugs and the sentence that relates them. The SQLite structure for the DDI database will be as described in Table 5.10.

```
−<sentence id="DrugDDI.d53.s4" origId="s4" text="In vitro studies also indicate that celecoxib, although
 not a substrate, is an inhibitor of cytochrome P450 2D6.">
   <entity id="DrugDDI.d53.s4.e0" origId="s4.p45" charOffset="36-45" type="drug" text="celecoxib"/>
   <entity id="DrugDDI.d53.s4.e1" origId="s4.p50" charOffset="92-111" type="drug" text="cytochrome
   P450 2D6"/>
   <pair id="DrugDDI.d53.s4.p0" e1="DrugDDI.d53.s4.e0" e2="DrugDDI.d53.s4.e1" interaction="true"/>
 </sentence>
```

Figure 5.4: XML for a DrugDDI Corpus example

Column	Description
Drug1	Drug1 name
Drug2	Drug2 name
Interaction sentence	The sentence that explains the interaction

Table 5.10: Drug DDI SQLite Database schema

The database originally contains approximately 3000 drug interactions and a total of 4900 chemical substances.

Since the source for this database is created based on InformationExtraction from medical documents, we have some entries with ambiguous or irrelevant information. For example:

Drug1 - antihypertensive drugs

Drug2 - hydralazine

Sentence - When other potent parental antihypertensive drugs, such as diazoxide, are used in combination with hydralazine, patients should be continuously observed for several hours for any excessive fall in blood pressure.

This interaction is not a direct interaction between substances. However it represents more than 1 interaction, the interaction between hydrazine and all the drugs that belong to the antihypertensive drug family. In order to solve this ambiguity problem we have to implement a more complex solution for the XML interpretation.

Other cases have irrelevant information with entries like "drugs" or "medication" that will not help us to get any interaction.

5.2.1 DDI Database Improvement

To create a more specific and accurate database we will use again the UMLS database to solve cases like the one previously presented. It is possible for us to get all the substances of a specific drug family. With this we can multiply ambiguous entries into chemical substance - chemical substance interactions. To do this we will use the queries such as the ones show in Listings 5.2 and 5.3.

```
1 Select CUI, LAT, ISPREF, STR
```

```
From MRCONSO
Where STR ="Antihypertensive drugs" and ISPREF ="Y";
+----------+-----+--------+-----------------------+
| CUI      | LAT | ISPREF | STR                   |
+----------+-----+--------+-----------------------+
| C0003364 | ENG | Y      | Antihypertensive Drugs |
| C0003364 | ENG | Y      | antihypertensive drugs |
+----------+-----+--------+-----------------------+
2 rows in set (0.01 sec)
```

Listing 5.2: Concept query

The process represented by Listing 5.2 will get us the identifiers available for the given String entry. We can use it in the next stage to get the relations between the other concept atoms or translations. The field "ISPREF" refers to preferential terms in use.

```
Select CUI1, CUI2, REL, RELA
From MRREL
Where REL="RB" and CUI2 ="C0003364" and RELA ="inverse_isa" and CUI1!="";

AND

select CUI1, CUI2, REL, RELA
From MRREL
Where CUI2 ="C0003364" and RELA ="mapped_from" AND SAB ="MSH";
+----------+----------+-----+-------------+
| CUI1     | CUI2     | REL | RELA        |
+----------+----------+-----+-------------+
| C0004975 | C0003364 | RB  | inverse_isa |
...
| C2743337 | C0003364 | RB  | inverse_isa |
| C2936669 | C0003364 | RB  | inverse_isa |
+----------+----------+-----+-------------+
120 rows in set (0.04 sec)
```

Listing 5.3: Relation query

As we can see there are two queries in this step represented by Listing 5.3. Only one of them will return results according to the type of relation the substances have with the entry family or source of information. This happens because UMLS gathers information from several sources and there may be different "RELA" fields that still identify a relation between concepts.

Then we will take all the different CUI's to check for relations.

```
Select CUI, LAT, ISPREF, STR
From MRCONSO
Where CUI="C2743337" and ISPREF ="Y";
```

```
+----------+-----+--------+-------------------------------------------------+
| CUI      | LAT | ISPREF | STR                                             |
+----------+-----+--------+-------------------------------------------------+
| C2743337 | ENG | Y      | clonidine, chlorthalidone drug combination      |
| C2743337 | ENG | Y      | clonidine - chlorthalidone - drug combination   |
+----------+-----+--------+-------------------------------------------------+
2 rows in set (0.03 sec)
```

Listing 5.4: CUI to concept name

The results returned by the query in Listing 5.4 will be individually analysed since some of them may need some editing or should not be considered. Listing 5.5 represents a set of possible names that can result for the execution of the described process.

```
1 - 5-(2-chlorophenyl)-7-fluoro-1,2-dihydro-8-methoxy-3-methylpyrazol(3,4b)(1,4)
    benzodiazepine
2 - Alprazolam - chemical (substance)
3 - Bromazepam (product)
4 - Methaminodiazepoxide
5 - Monopotassium, Clorazepate
6 - N-Descyclopropylmethylprazepam
7 - N-Descyclopropylmethylprazepam
8 - N-Descyclopropylmethylprazepam
9 - Diazepam [anxiolytic] (product)
10 - acetaminophen + ibuprofen + codeine
```

Listing 5.5: CUI to concept name

In the example result set presented by Listing 5.5 we have some examples of results that need editing. Elements that are presented in their chemical form as the result number 1 do not need to be taken into account, since we will only analyse their chemical substance or brand names. Elements 2, 3, 5 and 9 contain extra information that is irrelevant to our purpose, so they will be edited. We can also check duplicated elements like 6, 7 and 8. The result number 10 contains multiple entries as one String. These entries must be separated and verified individually.

To analyse and edit the String entries we used he following regular expression (RE) defined in JAVA:

```
(\w\-?\s?\w?)+\s?(\(\w*\))?(\[\w\])?((\[\w*\])\s?(\(\w*\)))?
```

This will limit the range of results to the ones we want and filter the extra unwanted information. We only limited the more recurrent cases, since it would be impossible to analyse all the possible scenarios. This RE will limit the entries to strings with simple names or string with both normal and straight parenthesis as the ones in line 3, 2 or 9 where we can easily extract the concept that we want.

Note that as we receive the query result for the CUI related with the drug family we will check the row with most occurrences using the STR field for the comparison. That STR name will be the one used to store in the SQLite Drug-Drug interactions database.

We still can limit the languages to English since our interactions database has that language. Using this method we can add several entries to our database based on previous information extracted from the XML files.

It is important to point that an interaction between Drug1 and Drug2 is the same as an interaction between Drug2 and Drug1. So, in the query process we must check for both scenarios, since there is nothing that ensuring that the database contains both pairs Drug1-Drug2 and Drug2-Drug1.

To check the accuracy of our drug-drug interactions database we checked for some known listed drug interactions. This test was done later on and is presented in the Chapter 6.

At this point, the information we will need to gather is ready to use. Figure 5.5 presents the work flow for drug-drug interactions overall evaluation process.

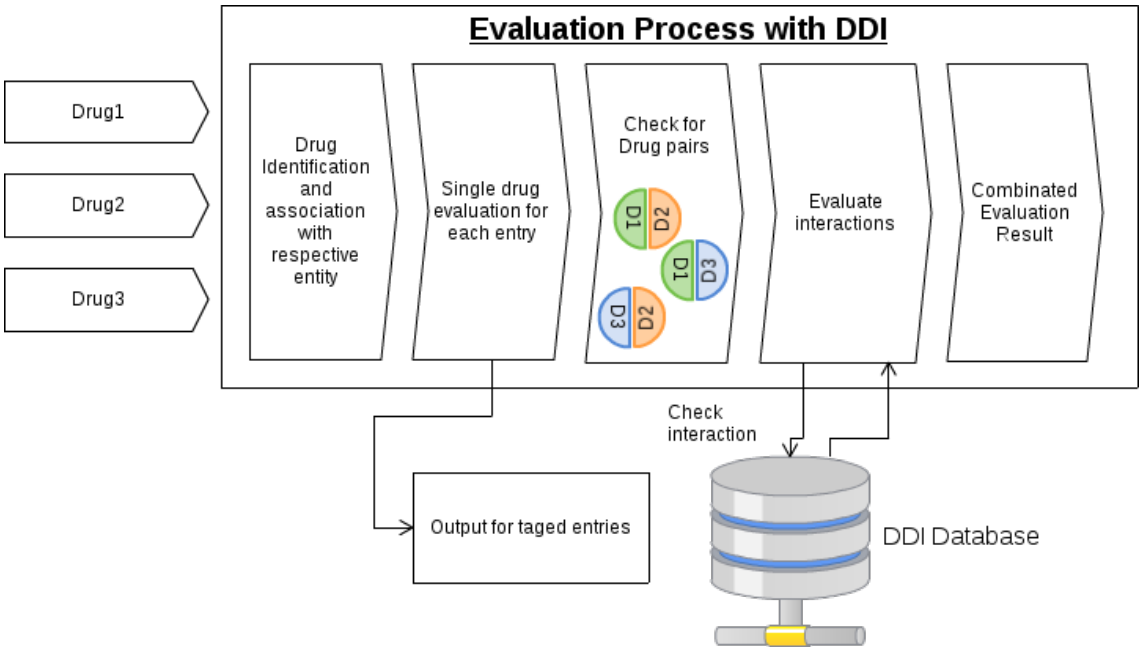


Figure 5.5: Overall evaluation process

The work flow starts by evaluating the entries and linking them to the respective drug concepts. The concepts will allow us to get the entity information we need. The second step represents all the work that is done in the single drug evaluation process. For detailed information on that process see Figure 5.1. The next step is where we start the real work on the drug interactions.

We will verify all the combinations of 2 drugs possible with the entries we have. All those combinations will be checked for interactions using the DDI database. The interactions will be verified in the given order and in inverse order to assure maximum matching. To improve matching and since the original source of information has a relatively small number of Database entries we

just load all the entries and match them with the system entries using Levenshtein Distance method with a NED value of 0,85.

As an additional step we can translate the entries and recheck if the original drug is not found. The possible combination side effects will be presented to the user if any entry was found. The data will contain the given drugs with the name of the database entries and the description associated with them regarding the interaction description.

5.3 Project Architecture

It is important from the developer's point of view to know the project layout and Java architecture, so here we will explain superficially the project architecture and function of the main parts.

5.3.1 SQLite DrugDB Manager

This class will be dealing with all SQLite drug database interactions. It works as a handler to help the other parts of the project, facilitating the communication with the databases. Classes like Drug Interaction Manager, MySQL Manager and Side Effects Manager will use this as a gateway to access, modify or add data.

5.3.2 Drug Interaction Manager

As the name discloses, this class is responsible for all operations related to the drug-drug Interactions. It will use all the information presented in the files from the DrugDDI Corpus, organize it in a structure and pass it to the DrugDB manager to be stored.

5.3.3 Side Effects TSV Manager

A class similar to the Drug Interaction Manager. It will access and transform the information available in the SIDER database files. The original download format of the SIDER datbase is as 3 TSV files that will be read by this class. In order to get the relations between drugs it will cross references between the files to get drug names, side effects and CUI's that are available on the records. When the system gets all the information organized it will pass it to the Drug DB manager to be stored and ready for usage.

5.3.4 MySQL Manager

This is an essential part of the project that will be in charge of all the communication with the UMLS database. All the queries here will be only of the access type since we have no interest in edit or add information to the UMLS dataset.

5.3.5 Project Manager

The Main class on the project. Here we can access all the information available that was organized before in the SQLite Databases. A debug interface represented in Figure 5.6 was created to access all the parts of the project individually that allow for the re-population of the ontology, queries to the SQLite databases, queries to the UMLS database and queries on the ontology itself.

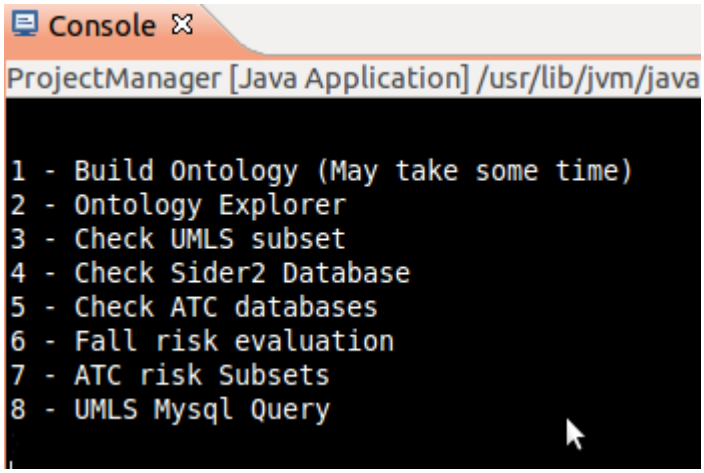


Figure 5.6: System debug menu

5.4 Output Result

We intend to have an available web service that will allow the users to query the system and retrieve a result. The system output allow to user to check the possible consequences of the combination of drugs being used. We can either query for a simple drug evaluation or query for combinations/interactions results. The system is able to receive an array of multiple drugs and evaluate them both as individuals and as a group for their interactions.

To demonstrate the system output we will use an entry of some risky drugs, combinations, tradenames or chemical substances, and some irrelevant drugs to our goal. We will use *Xanax*(tradename) as a fall related drug, *Aspirine*(tradename) as an irrelevant entry, *warfarin* and *ciprofloxacin* to check for drug-drug interactions.

5.4.1 Single drug output

Here we will review the part of the result relative to the individual entry drugs evaluation. Since the result output on the program will, for debug purposes, present the sources and information flow we will only show the output relative to relevant conclusions and data. The Side effects list will be as well limited and can be consulted in full in the appendix A.1.

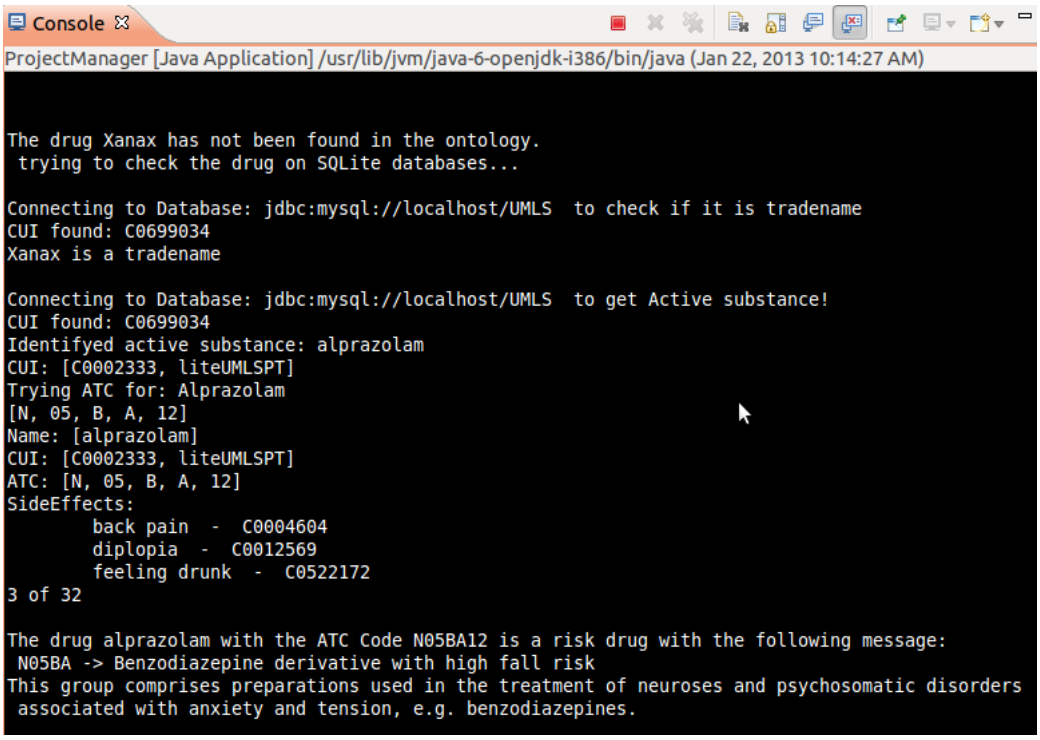


Figure 5.7: Xanax system output

As we can see in Figure 5.7, *Xanax* was correctly evaluated as a trade name of the active ingredient/substance alprozam. Then it was identified as a member of the benzodiazepines which are one of the main classes related to falls.

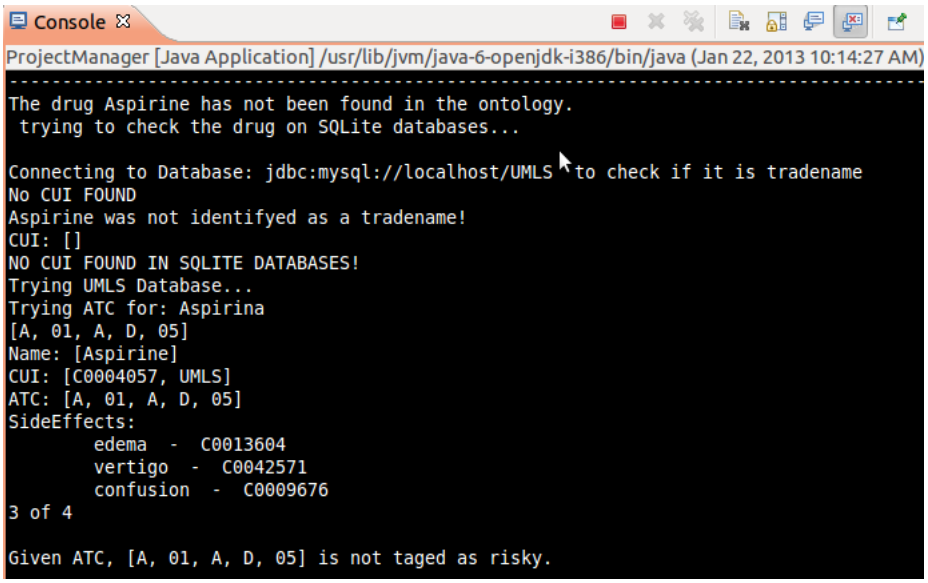


Figure 5.8: Aspirine system output

The second entry in Figure 5.8, *Aspirine*, was equally identified and evaluated by the system. There are some minor situational side effects, but the overall evaluation points to no associated risk.

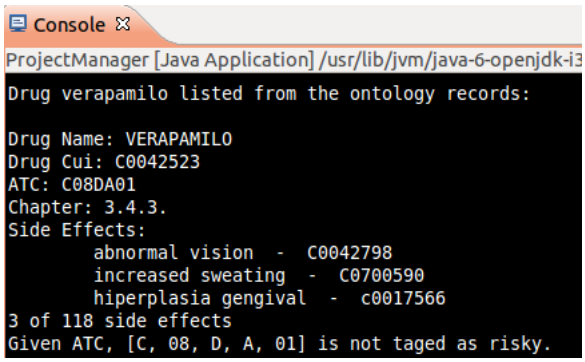


Figure 5.9: Verapamilo system output

The concept for *verapamilo* in Figure 5.9 was found on the ontology population, so the evaluation was done using that data previously updated with information from the ATC gathered data and a few added side-effects. The original ontology already had several side effects that should be filtered to the ones related with falls as it is done with data outside the ontology. However, and aside the vast number of side-effects that can be related to verampil, the drug is not part of a high fall risk drug group.

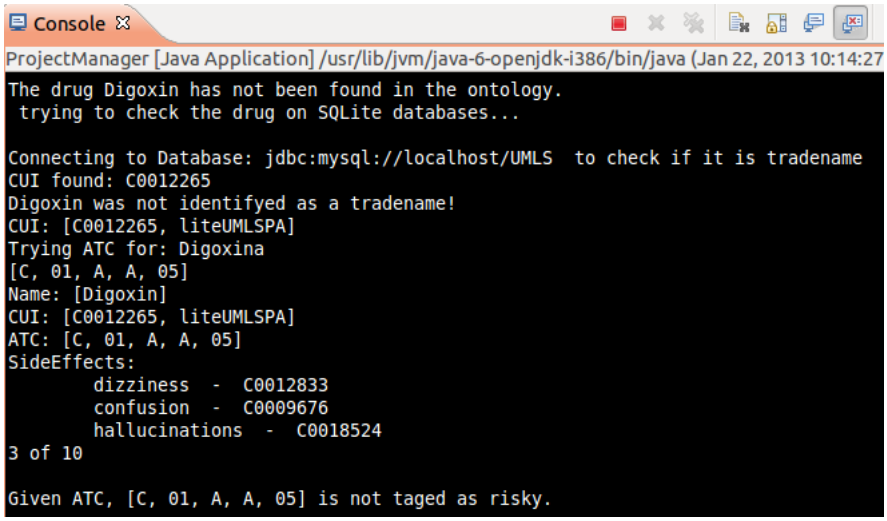


Figure 5.10: Digoxin system output

As we can see in Figure 5.10, *digoxin* on its own is not considered risky for falls either. It also presents some side effects that may increase the fall risk associated, however, these side effects were not retrieved using the original ontology population, but the updated side effects originally from SIDER source database.

5.4.2 Drug-Drug Interaction output

For the 4 entry and since the order is irrelevant there are 6 possible combinations. The system will identify those 6 combinations and then query the interaction database for any possible entry. The result will be presented with the names of the 2 drugs and the query result for sentences that relates both drug concepts.

Once again, the entire result set for this example is in the appendix A.2 and here we will only show in Figure 5.11, one of the interactions and explain some details that might not be obvious to a non specialist in the field.

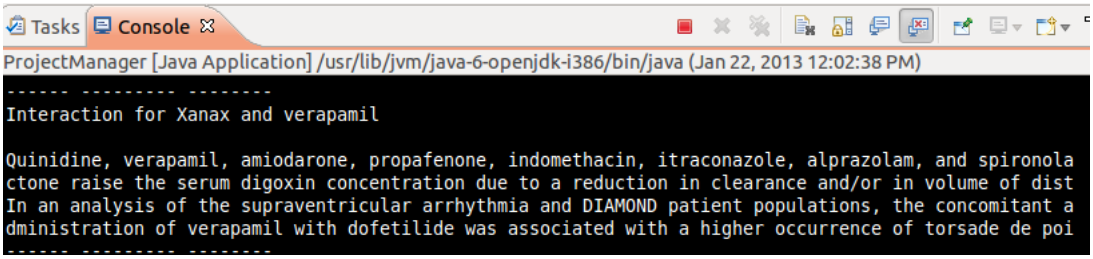


Figure 5.11: Interaction between Xanax and Verapamil

When analysing the first sentence we will not see any direct relation between Xanax and Verapamil, since Xanax is not even in the sentence. However and due to the database disambiguation process that we mentioned in Section 5.2.1 we added information relative to terms that represent a family or a group of substances. In this case, *Xanax* is a tradename of *Alprozam* that is mentioned in the text. There are several cases where this happens during the identification of drug interactions.

Chapter 6

Result Analysis

This analysis will be conducted in order to study the process of re-population/update of our data model and to evaluate the accuracy of the system response. To do this we will at first check the entries of the original ontology population that we were able to identify and update with new fall related information. To check if this process was conducted with loss of information due to miss match between the original ontology population and our sources we will do a match evaluation.

As the second phase of the evaluation we will use available lists of drug-drug interactions and fall risk medication to access the veracity of the result given by this thesis system. This is an overall system evaluation and will use verified information in the field based news/articles/studies.

6.1 Update of the original data model population

Here we will conduct a simple test regarding the effectiveness of linkage between the various information sources we will need. The objective is to use the original ontology population and conduct a study that will retrieve a matching ratio and complete the updated data model with new information. Ideally all the entries of the original population would be found, however, that result is unlikely since the original population contains some entries that may not be considered exactly medical drugs or simply are not found in our databases. Those cases will be explained in detail later on this Chapter. Anyhow, the system sources regarding the chemical substances and entities are extensive and reliable, so we have good matching perspectives and consequently the update of the ontology model should be complete.

To conduct the tests we developed an extra class in the project architecture called MatchingTestClass that is responsible for all the matching and checking tests.

6.1.1 Concept identifying process

In order to get a great part of the information we need a link between the original population and the UMLS Database. To check if that link is done properly we developed this test/verification.

The UMLS is the most complete database we have at our disposal, so we hope to achieve a high match ratio for this verification. Firstly we decided to match the concepts using only the SQLite subsets. The process will prioritize the SQLite PT subset, then the SPA subset and at last the ENG subset. We obtained the results shown in Table 6.1.

Subset	Matches	Total Remaining	Match Ratio
POR	530	1359	39%
SPA	327	829	39,44%
ENG	42	502	8,36%
Total	899	1359	66,15%

Table 6.1: Concept matching results

Note that the verification to the Spanish subset does not contain the concepts already found in the Portuguese search. The same thing happens in the English subset.

Using only subset information we were able to link 66,15% of the total entities available in the ontology. However, to increase this number we will submit the remaining entries to a search using UMLS MySQL complete database. The results retrived from this study are presented in Table 6.2.

Database	Matches	Total remaining	Match Ratio
UMLS	97	405	24%
Absolute total	966	1359	73%

Table 6.2: UMLS concept matching results

The majority of non-match entries are due to no CUI entry because they are not considered actual chemical substances, no translation available or minor name differences in the UMLS database where we can not apply the Levenshtein Distance method in the query.

6.1.2 ATC linkage process

This evaluation will be made according to the steps used in the re-population algorithm. We will check all the different entries in the original ontology set. This means that we have to divide the composed entries and check for duplicates. After the definitive list of entries is ready we will check for the translations using UMLS subsets to match with the ATC sources. If a direct match is not possible with the translation and ATC entries we will try a Levenshtein Distance match with a NED value of 0,85.

The direct matching algorithm without translation obtained the results shown in Table 6.3. The results are base on the original ontology population with a total of 1359 different concepts.

Since this first part of the ATC matching analysis only uses direct matching the low ratio may be considered normal due to the limitation in the databases to 2 languages and no checking in the translations. Using the SQLite database to check for translation improved result were achieved. The data on those results are presented in Table 6.4.

SQLite Source	Matches	Match Ratio
ATC ENG	185	13,6%
ATC GER	265	19,5%
Both	182	13,5%
ATC Total	268	19,7%

Table 6.3: ATC direct match results

Source	Matches	Total translated	Match Ratio
ATC ENG	676	857	78,87%
ATC GER	520	857	60,68%

Table 6.4: ATC direct match results with translation

A total of 857 entities refers to the entries that had CUI and we could translate. As it is proved by the results the matching ratio increased immensely when we search for concepts using the same input language. Nevertheless, these results can still be improved and we have to find a method to apply to the concepts with no translation.

In many translations the only thing that will differ from one language to another is the last character, accentuation or a different character. This problem can be easily solved. To do this we conducted the following matching improved with Levenshtein distance in the concepts without CUI and in the concepts that still had no match, a total of 574 entries, as presented in Table 6.5.

Source	Matches	Total Translated	Match Ratio
ATC ENG	169	574	29,4%
ATC GER	84	574	14,5%

Table 6.5: ATC improved match results

Since all the results we found in ATC_GER are also found in ATC_ENG we do not have any overlap and with these last improvements we reached a total and final match ratio of 70,2% verified in the algorithm and explained by the equations 6.1 and 6.2.

$$TotalPopulation - (Remaining - newMatch) = TotalMatch \tag{6.1}$$

Where we apply the values:

$$1359 - (574 - 169) = 954 \tag{6.2}$$

The remaining 30% drugs will have to be evaluated based only on the side effects associated with them. After all the efforts to improve the matching result we think that the only way to improve even more the data association is by using a more complete ATC source.

6.1.3 Side effects update

The evaluation of side effects entries will be done according to the same principles used on the ATC evaluation. We will check for direct matches between the names in the original data model population and if with that match it is not possible we will try a match using Levenshtein Distance with the same NED value of 0,85.

The direct results of this methods are shown in Table 6.6.

Operation	Result	Total
Matched	590	1359
Not matched	768	1359
Not Translated	460	1359

Table 6.6: Side Effects matching results

There were some entries where even without translation we managed to have a match using the Levenshtein Distance. However, there are still several concepts without translation. We decided to conduct a more extensive test to find out why this was happening.

On the total of 768 entries that were not found, 371 were not even accessible via UMLS database, the most complete database we had access to. The reason for this is simple after checking some of the entries that were not found:

- brometo declidinio
- pidolato de calcio
- 0.12 mg/24 hnuvaring1 unid
- Tintura de Bendoim
- Sais Minerais
- vacina contra hepatite C
- vacina contra a difteria, o tetano,a tosse convulsa e o haemophilustipo b

Some concepts were not identified because they were wrongly recorded in the data model during the extraction process like *brometo declidinio*, *0.12 mg/24 hnuvaring1 unid;* or *cloreto depotassio*. Some of these cases could be solved with Levenshtein Distance, but if the source for the concept was the UMLS MySQL database we can not apply a direct query using that method.

Some other concepts may be easily recognisable by a professional, but they are on record with a non-scientific or non appropriate nomenclature, therefore the entry in the UMLS dataset may not exist. Examples like *Sais Minerais* and *Tintura de Bendoim* can reflect those cases.

Another reason for those mismatches is the vague concept listed in the entry name. Examples of these cases like *vacina contra hepatite C* or *vacina contra a difteria, o tetano,a tosse convulsa e o haemophilustipo b*. The list of all these 371 concepts is available in the Appendix A.3.

The remaining 397 entries that had translation but still had no match on the side effects database simply were not listed in the source.

Most of the errors presented here are normal since the original ontology population was all created using Information extraction. However, with the translation and concept identifying system we used we were able to minimize the mismatches in the model population.

6.1.4 Drug interactions

As mentioned before, in this part of the system we will add a direct interaction relation between the concepts if they both exist in the ontology. Otherwise, if only one of the drugs exists in the original model population we will add a Data property identifying the name of the interacting drug and the respective CUI. If none of the drugs exists in the dataset we will not add anything.

In this test we will evaluate the concept linkage between the original population ontology and our improved interaction database. Table 6.7 present the results obtained.

Operation	Result	Total
Matched	479	1359
Not matched	880	1359

Table 6.7: Ontology concept linkage to the DDI database

As we can see, based on these results, the matching ratio we achieved is not optimal. This fact can be justified since the ontology population and the Drug DDI population are both extracted and created using Information Extraction from documents and are not optimal for this kind of usage. The accumulation of possible extraction errors from both parts causes a decrease in the matching ratio. We tried to improve this match ratio during the Interaction database improvement process by expanding the concepts and by using approximated string matches (Levenshtein Distance), however, in order to achieve a more accurate result the best approach is the utilization of another source for drug-drug interactions.

6.2 Result accuracy evaluation

This section's objective is to evaluate the overall results obtained by the system. It intends to check if the retrieved results are in fact correct. To do this we will use a set of drugs listed by an official document and manually checked out by healthcare professionalsA.4. The existing drugs on the document will be fed to the system and hopefully it will identify them and their drug family as risk related.

The other part of this evaluation will be about the interactions. To assess the accuracy here we will use again verified data and compare the verified results to this system's results.

6.2.1 Single drug

To do this evaluation we used a list of drugs that are considered risky for falls from UNC-Chapel Hill news [Etc08]. This list contains a total of 93 drug entries with trade names and chemical substances. It can be consulted in the appendix A.4 of this document.

A total of 83 Drug entries were correctly identified as risky based on the ATC match. The 10 remaining entries were not identified as risky for various reasons. Table 6.8 shows the details on the remaining substances and we will suggest how the results could be improved.

Substance	CUI	ATC	Side-effects
Oxymorphone	✓	-	-
Levorphanol	✓	-	-
Butabarbital	✓	-	-
Clidinium-chlordiazepoxide	-	-	-
Codeine	✓	R05DA04	✓
Digoxin	✓	C01AA05	✓
Disopyramide	✓	C01BA03	-
Divalproex sodium	✓	-	-
Hydrocodone	✓	R05DA03	-
Duloxetine	✓	G04BX18	✓

Table 6.8: Accuracy results

As we can see, the majority of the drug entries were identified and related to the respective concepts in the UMLS database. However, only half of the concepts were successfully connected to the ATC entry. This problem alone represents a big problem in the evaluation process, however, we can not do anything about these specific cases since we are limited to the ATC database information available to us.

Another problem verified in 4 out of 10 concepts was the identification of the ATC as not risky. This happens because these specific drugs represent exceptions. The drug family/type of this drugs are not linked to any increase in the fall ratio. However, these specific drugs are, so, what we can do is add the CUI relative to this specific substance as risky. Adding the substance family or sub-family to the risky ATCs would be a mistake once it would falsely assess several substances as risky.

The overall result of this test point to a correct identification of 83 out of 93 substances as risky. This represents a success ratio of 89,25%. A concept identification ratio of 98,89% leaving only *Clidinium-chlordiazepoxide* unlinked and an ATC identification ratio of 94,62%.

6.2.2 Interaction study

Here we will need to develop a similar method as the one used in the fall evaluation for Single drug. To do this the first step we have to accomplish is to find a viable list of interactions in order to check the quantity and quality of our database information. As we mentioned before, in this case our source dataset was not ideal due to the quantity of information and generation method.

Anyhow we expect it to have a decent performance since we worked on the database improvement. We did not add any data manually, so all the extra information was created using inference. The process is explained in Section 5.2.1.

For result comparison purposes we used a list of 10 drug interaction for long term patients by SeniorJournal.com [1].

- Warfarin — NSAIDs [2]
- Warfarin — Sulfa drugs
- Warfarin — Macrolides
- Warfarin — Quinolones [3]
- Warfarin — Phenytoin
- ACE inhibitors — Potassium supplements
- ACE inhibitors — Spironolactone
- Digoxin — Amiodarone
- Digoxin — Verapamil
- Theophylline — Quinolones [3]

As we can see this scenario is more complex than a simple database query since we have first to check for non chemical substance concepts like *Quinolones* or *NSAIDs*.

We used the search algorithm and obtained the Table 6.9:

Drug1	Drug2	Interaction entries
Warfarin	NSAIDs	27
Warfarin	Sulfa drugs	0
Warfarin	Macrolides	0
Warfarin	Quinolones	11
Warfarin	Phenytoin	20
ACE inhibitors	Potassium	5
ACE inhibitors	Spironolactone	0
Digoxin	Amiodarone	1
Digoxin	Verapamil	2
Theophylline	Quinolones	5

Table 6.9: Interactions Results

As verified, only 3 of the 10 entries did not receive any query result. The entries were checked to see if they are a generalist concept and expanded into the concept components. If they still

[1] http://seniorjournal.com/NEWS/Eldercare/4-12-14TenDrugs.htm
[2] NSAID class does not include COX-2 inhibitors
[3] Quinolones that interact include: ciprofloxacin, enoxacin, norfloxacin and ofloxacin

have no result that means that our source did not contain any relevant information to the given combination. To check this we searched manually for drug interactions between substances of the families that were not matched and got no results as expected.

Chapter 7

Conclusion

After all the work developed during this project it is important to review the process and identify flaws and mistakes in order to improve future work. In this section we will overview the last months of work and point some problems that were identified and the perspectives of solution to those problems. We will also talk about imperfections of this tool that can be corrected and improved in the future.

7.1 Objectives overview

In this section we will evaluate the fulfilment of the proposed objectives. The project turned out to be slightly different from what i was initially anticipating with a lot of work regarding concepts and information linkage. The knowledge representation model was indeed a very important part of the system, responsible for the storage and organization of all information.

The main objective of this dissertation is the evaluation of the impact of single and multi drug usage on falls in the elderly population. To do this several objectives were outlined in Chapter 1.

We started by developing an updated model of the existing knowledge representation model that would be capable of representing the medicine and the development of an algorithm capable of reaching conclusions on the side effects of multiple medications. The system should be capable of identifying drug entries representing chemical substances or drug trade names.

Later on the entries should be related to the respective concept or entity in the data model so they can be analysed properly. If the entry represents a trade name the system should also be able to retrieve the respective chemical substance to evaluation purposes. After having the concept name and substances identified the system should access the related side effects, drug class classification and make a fall risk evaluation for a single drug or combination of drugs.

The currently updated ontology model is capable of representing all the concepts related to the medication needed for our algorithm in order to evaluate the fall risk. It still contains data from the original model that is not used during the evaluation, however it makes no sense to remove

those parts of the ontology since it will allow the usage of the ontology for other purposes and is no obstacle to our objective.

The ontology population update process was as well successfully implemented. Using the external sources, the system checks and identifies the concepts, linking them to the entries existing in the external sources. Once this linkage is done, the process continues and new information is added to the model.

The final objective, the evaluation of fall risk, was as well fulfilled with good results and using all the data from the previously named objectives.

7.2 Conclusion

The usage of medical electronic systems, digital information records and decision support systems are becoming in the current days more and more frequent. They represent a way to improve the health system with a small cost. This dissertation intends to use the available digital records about the drugs and create a tool that can help using and analysing that information to identify possible falls risks for single drug and drug combinations in elder persons.

Overall, the result we achieved in this project were a success. Using known sets of drugs that have already been proven to have a relation to the falls in older people, we managed to test our system qualitatively and quantitatively. The success ratio in this tests were 90% for single drugs and 70% for drug-drug interactions.

In conclusion, the development of a system with an entire work flow, where we have drug identification, concept association and evaluation, was executed.

The system can be used in collaboration with several projects aiming for the study and prevention of falls. The basic interaction process with external medical sources can also be used to achieve other search or evaluation processes in the medical/pharmaceutical field.

7.3 Improvements and future work

After gaining some experience in the field and knowledge of the problem and the project there are a few things that should have been done differently. First of all the usage of the UMLS language subsets. It is in fact a faster query process for the translation process, even so, the gain is not substantial, the work time used to create the databases was significant and we still need to use the UMLS MySQL database to do queries that are not related to the translation process. The subsets were created in a first stage of the system development where we did not predict the need for some other queries to the UMLS database.

A substantial improvement to the project would be the introduction of a more complete and accurate interactions database, since the current one was developed out of a corpus with a completely different objective than the one we have here.

Conclusion

The matching process was improved regularly during the development of the project, however, there still is room for more improvements by matching accuracy and quantity or even by performance. This last factor was taken into account and greatly improved during the project but not the focus of our attention since it represents a secondary goal, being the priority the effectiveness of the system.

Associação Fraunhofer Portugal intends to use the work developed during this thesis and improve it to be used as a part of other fall related projects currently under development.

Appendix A

Appendix

A.1 Single drug evaluation Output

```

The drug Xanax has not been found in the ontology.
 trying to check the drug on SQLite databases...

Connecting to Database: jdbc:mysql://localhost/UMLS  to check if it is tradename
CUI found: C0699034
Xanax is a tradename

Connecting to Database: jdbc:mysql://localhost/UMLS  to get Active substance!
CUI found: C0699034
Identifyed active substance: alprazolam
CUI: [C0002333, liteUMLSPT]
Trying ATC for: Alprazolam
[N, 05, B, A, 12]
Name: [alprazolam]
CUI: [C0002333, liteUMLSPT]
ATC: [N, 05, B, A, 12]
SideEffects:
back pain  -  C0004604
diplopia  -  C0012569
feeling drunk  -  C0522172
confusion  -  C0009676
incontinence  -  C0021167
anorexia  -  C0003123
dystonia  -  C0393593
incoordination  -  C0520966
blurred vision  -  C0344232
somnolence  -  C2830004
edema  -  C0013604
ataxia  -  C0004134
```

```
weakness  -  C1883552
asthenia  -  C0004093
vertigo  -  C0042571
muscle stiffness  -  C0221170
tremor  -  C0040822
muscle weakness  -  C0151786
dyskinesia  -  C0013384
Cognitive Disorder  -  C0009241
peripheral edema  -  C0085649
cramps  -  C0026821
drowsiness  -  C0013144
hallucinations  -  C0018524
euphoria  -  C0235146
abdominal pain  -  C0000737
Mental impairment  -  C1306341
insomnia  -  C0917801
stiffness  -  C0427008
arthralgia  -  C0003862
hypotension  -  C0020649
headache  -  C0018681

The drug alprazolam with the ATC Code N05BA12 is a risk drug with the following message:
 N05BA -> Benzodiazepine derivative with high fall risk
This group comprises preparations used in the treatment of neuroses and psychosomatic disorde

--------------------------------------------------------------------------------
The drug Aspirine has not been found in the ontology.
 trying to check the drug on SQLite databases...

Connecting to Database: jdbc:mysql://localhost/UMLS  to check if it is tradename
No CUI FOUND
Aspirine was not identifyed as a tradename!
CUI: []
NO CUI FOUND IN SQLITE DATABASES!
Trying UMLS Database...
Trying ATC for: Aspirina
[A, 01, A, D, 05]
Name: [Aspirine]
CUI: [C0004057, UMLS]
ATC: [A, 01, A, D, 05]
SideEffects:
edema  -  C0013604
vertigo  -  C0042571
confusion  -  C0009676
drowsiness  -  C0013144
```

```
                                                                                                    44
Given ATC, [A, 01, A, D, 05] is not taged as risky.

2 ------------------------------------------------------------------------------------------------
Drug verapamilo listed from the ontology records:
4
Drug Name: VERAPAMILO
6 Drug Cui: C0042523
ATC: C08DA01
8 Chapter: 3.4.3.
Side Effects:
10 abnormal vision - C0042798
increased sweating - C0700590
12 hiperplasia gengival - c0017566
Symptoms of hepatitis - C0858814
14 flushing - C0016382
sinus arrest - C0178428
16 Sinusite - c0037199
paralisia intestinal - c0030446
18 paralysis - C0522224
Icterícia - c0022346
20 fraqueza muscular - c0151786
edema periférico - c0085649
22 dor torácica - c0008031
dor torácica - c0030193 c0004757
24 macula - C0332573
Fadiga - c0015672
26 hypertrophic subaortic stenosis - C0700053
dissociation - C0086168
28 extrapyramidal syndrome - C0015371
upper respiratory tract infection - C0041912
30 polaquiúria - c0042023
keratosis - C0022593
32 dores articulares - c0003862
palpitações - c0030252
34 síncope - c0039070
enlargement - C0020564
36 xerostomia pós-irradiação - c0043352 pós-irradiação
numbness of extremities - C0239375
38 desconforto abdominal - c0232487
sinais extrapiramidais - c0234133
40 rinite - c0035455
Vasculite - c0042384
42 Bloqueios auriculoventriculares - c0004245
ileus - C1258215
```

respiratory failure - C1145670
Urticária - c0042109
edema dos tornozelos - c0235439
somnolence - C2830004
faringite - c0031350
parestesias - c0030554
trémulo - c0392703
colestase - c0008370
quadriparesis - C0270790
hypertrophic cardiomyopathy - C0007194
Obstipação - c0009806
laryngospasm - C0023066
causar hipotensão ortostática associada a tonturas - c0085978 c0020651 c1608982 c0012833
hiperplasia - c0020507
mal-estar gastrintestinal - c0234215 gastrintestinal
seizures - C0036572
gastrointestinal distress - C0548823
eritema polimorfo - c0014742
desequilíbrios hidroelectrolíticos - c0281825 hidroelectrolíticos
Alopécia - c0002170
reacções de hipersensibilidade com choque - c0020517 c0036974
tratamento da disfunção eréctil - c0039796 c0242350
diplopia - c0012569
mialgias - c0231528
ginecomastia - c0018418
dermatoses hiperqueratósicas - c0037274 c0870082
complete atrioventricular block - C0151517
Vertigem de causa - c0042571 c0085978
maculopapular rash - C0423791
Edema ligeiros - c0013604 ligeiros
broncospasmo passageiro - c0006266 passageiro
edema pulmonar - c0034063
surgir ainda dispneia ainda - c1948041 c0013404 ainda
Hepatite B em doentes com infecção - c0019158 c0004690 c0030705 c0021311
weakness - C1883552
ECG abnormal - C0522055
ecchymosis - C0013491
angioedema - c0002994
casos de eritema com sintomas - c1053418 c0041834 c0683368
cardiomiopatia - c0878544
sinais de intoxicação incluem visão turva - c0220912 c0728899 c0344232
Astenia - c0004093
coldness - C0812387
prolactin increased - C0553731
claudication - C1456822

```
Dispepsia - c0013395
oligomenorreia - c0028949
hiperprolactinemia - c0020514
hypertensive - C0857121
Tratamento de sintomas depressivos - c0039796 c0086132
Alterações do sono - c0037317
tratamento de infecções respiratórias - c0039796 c0021311 respiratórias
tratamento de infecções respiratórias - c0039796 c0039796 c0035243
púrpura - c0034150
diárias - c0011991
confusão mental - c0009676
sulfonamida incluem erupção cutânea - c0038760 c0015230
dor associada ao enfarte do miocárdio - c0030193 c1608982 c0027051
gripe - c0021400
atrioventricular dissociation - C0004331
elevated liver enzymes - C0235996
acidente vascular cerebral - c0038454 c0238051
allergy aggravated - C0235893
casos descritos de galactorreia - c1053418 c0038545 c0235660
tinea cruris - c0040264 cruris
tinea cruris - C1384589
controlo das náuseas - c0243148 c0027497
prurido das alergias cutâneas - c0033774 c0020517 cutâneas
sudorese - c0038990
cãibras - c0026821
Cefaleias - c0018681
nistagmo - c0028738
numbness - C0028643
SNC incluíndo insónia - c0927232 c0917801
tratamento de doentes com infecção - c0039796 c0030705 c0021311
choque anafiláctico - c0002792
drowsiness - C0013144
tremor - c0040822
vómitos post-operatórios - c0042963 post-operatórios
Letargia - c0023380
dor ligeira - c0030193 ligeira
prevenção de espasmos - c0199176 c0037763
muscle fatigue - C0242979
IC congestiva - c0018801 congestiva
Given ATC, [C, 08, D, A, 01] is not taged as risky.
--------------------------------------------------------------------------------
The drug Digoxin has not been found in the ontology.
 trying to check the drug on SQLite databases...

Connecting to Database: jdbc:mysql://localhost/UMLS  to check if it is tradename
```

```
CUI found: C0012265
Digoxin was not identifyed as a tradename!
CUI: [C0012265, liteUMLSPA]
Trying ATC for: Digoxina
[C, 01, A, A, 05]
Name: [Digoxin]
CUI: [C0012265, liteUMLSPA]
ATC: [C, 01, A, A, 05]
SideEffects:
dizziness  -  C0012833
confusion  -  C0009676
hallucinations  -  C0018524
abdominal pain  -  C0000737
weakness  -  C1883552
visual disturbances  -  C0547030
psychosis  -  C0033975
anorexia  -  C0003123
bradycardia  -  C0428977
headache  -  C0018681

Given ATC, [C, 01, A, A, 05] is not taged as risky.

--------------------------------------------------------------------------------
```

A.2 Drug-Drug Interaction output

```
------ --------- --------
Interaction for Xanax and verapamil

Quinidine, verapamil, amiodarone, propafenone, indomethacin, itraconazole, alprazolam, and sp
In an analysis of the supraventricular arrhythmia and DIAMOND patient populations, the concor
------ --------- --------
Interaction for Xanax and Digoxin

Cyclosporine, Digoxin, Methotrexate Lodine, like other NSAIDs, through effects on renal prost
Digoxin Some calcium blockers may increase the concentration of digitalis preparations in the
Digoxin and verapamil use may be rarely associated with ventricular fibrillation when combine
------ --------- --------
Interaction for Aspirine and verapamil

Quinidine, verapamil, amiodarone, propafenone, indomethacin, itraconazole, alprazolam, and sp
In an analysis of the supraventricular arrhythmia and DIAMOND patient populations, the concor
------ --------- --------
Interaction for Aspirine and Digoxin
```

```
                                                                                                    42
Cyclosporine, Digoxin, Methotrexate Lodine, like other NSAIDs, through effects on renal
Digoxin Some calcium blockers may increase the concentration of digitalis preparations
Digoxin and verapamil use may be rarely associated with ventricular fibrillation when co
------ --------- --------
Interaction for verapamil and Digoxin

Cyclosporine, Digoxin, Methotrexate Lodine, like other NSAIDs, through effects on renal
Quinidine, verapamil, amiodarone, propafenone, indomethacin, itraconazole, alprazolam, a
Digoxin Some calcium blockers may increase the concentration of digitalis preparations
Digoxin and verapamil use may be rarely associated with ventricular fibrillation when co
```

A.3 Side Effects not matched

CUI FOUND:397 NOT CUI FOUND:371

- levofolinato de calcio
- brivudina
- brometo de ipratropio
- alfatocoferol
- vacina contra a hepatite b
- fenspirida
- oleo essencial de tomilho
- tintura de benjoim
- tintura deeucalipto
- maca reineta
- nicotinato de benzilo
- pirissudanol
- peginterferao alfa-2b
- prulifloxacina
- oxibuprocaina
- propinoxato
- multivitaminas
- sais minerais
- dalteparina sodica
- bromelaina
- brometo de ipratropio
- lutropina alfa
- acido hialuronico
- multivitaminas
- nimesulida
- vitaminas do complexo b

acidoascorbico
carbonato de di-hidroxidodealuminio e sodio
polistireno sulfonato de calcio
celulas vivas de levedura
oleode figado de tubarao
vacina contra o meningococo
oxido de zinco
dioxido de titanio
hamamelia
cloro-hexidina
interferao beta-1b
glicerofosfato demagnesio
messalazina
sulodexida
vacina contra a difteria e omeningococo
acetonido de fluocinolona
subgalhatode bismuto
clorotalidona
cafeina
aminoacidos
vacina viva contra o rotavirus
peroxido debenzoilo
rupatadina
vinpocetina
carmelose
clorodiazepoxido
sulbutiamina
acetilsalicilato de lisina
folitropina alfa
fosfato tricalcico
proglumetacina
salicilato demetilo e outras associacoes
acido borico
hidroxido de aluminio
carbomero
ebastina
aceglumato de deanol
mononitrato de isossorbida
galium molugo
propifenazona

cafeina
fondaparinux sodico
fluoreto de sodio
fluorofosfato de sodio
vacina contra a hepatite a e ahepatite b
nonivamida
benzilpenicilina potassica
polisulfato sodico de pentosano
azuleno
vacina contra a difteria, o tetanoe a tosse convulsa
sais minerais e outrasassociacoes
acetato de calcio
acido ibandronico
levodropropizina
beladona (alcaloides)
cloreto decetilpiridinio
alcool diclorobenzilico
“heparinoide”
aceponato de metilprednisolona
para-diclorobenzeno
acido cromoglicico
etofenamato
hidroxido de magnesio
vacina contra o haemophilus tipo b
acetato de calcio
carbonato demagnesio
anacinra
cloreto de zinco
mirtecaina
butilescopolamina
peroxido de hidrogenio
carbonato de di-hidroxidodealuminio e sodio
clonixina
indobufeno
acido nicotinico
interferao gama-1b
alcatrao mineral
piritiona dezinco e outras associacoes
imunoglobulina humana contra ocitomegalovirus
cloreto de sodio

cloreto de potassio
glucoronamida
risedronato de sodio
plantago ovata (sementes)
hidrosmina
macrogol e outras associacoes
vacina pneumococica poliosidica
acetato de eslicarbazepina
deflazacorte
tetrizolina
acido lactico
lecitina de soja
oleo de soja
clorotalidona
mesilato de di-hidroergocriptina
cefodizima sodica
iodeto de tibezonio
lisado de escherichia coli
vitaminas do complexo b
trepibutona
tolterrodina
salicilato demetilo
oxibuprocaina
proteinosuccinilato de ferro
butamirato
aliscireno
cobamamida
brometo de pinaverio
promestrieno
picetoprofeno
oxolamina
sevelamero
oxido de zinco
rilmenidina
lisados polibacterianos
oleo essencial de eucalipto
oleoessencial de terebintina
acetonido de fluocinolona
iodopovidona
difluocortolona

brometo de tiotropio
cinchocaina
tiocolquicosido
hidrogenofosfato de calcio
insulina aspartico
tri-hexifenidilo
cleboprida
palonossetrom
citrato de potassio
lactobacillusacidophilus
cetrorrelix
diacereina
acido lactico
sabalserulata
citrato de potassio
peginterferao alfa-2a
mesilato de di-hidroergotamina
piritiona zinco
picossulfato de sodio
brometo de domifeno
tribenosido
ciamemazina
tromantadina
buserrelina
cloreto dedequalinio
factor viii da coagulacao humana
peroxido de benzoilo
interferao alfa-2a
mesilato de di-hidroergocristina
dicloverina
fosfato dissodico
fosfatomonossodico
pangamato de calcio
hidroxido de aluminio
hidroxidode magnesio
tansulosina
peroxido de benzoilo
fenolftaleina e outrasassociacoes
laurilsulfoacetato de sodio
nadifloxacina

ustecinumab
valerato deestradiol
carboximaltose ferrica
cromocarbo dietilamina
benzilpenicilina procainica
dapoxetina
acidosalicilico
mirtecaina
cloreto de sodio
flupirtina
hipericao
colecalciferole
sais minerais
fluoreto de sodio
agomelatina
oleo de soja
etofibrato
ruscogenina
nilvadipina
brometo de distigmina
sulfadiazina prata
”heparinoide”
insulina isofanica
ictiol
oleo de soja
di-hexazina407
cloreto de zinco
cloro-hexidina
cloreto de trospio
clorofenoxamina-
hidrolisado cerebral de porco
ictiol
lecitina desoja
ispagula (tegumento)
ispagula(semente)
oxitriptano
pidolato de magnesio
dectafluor
fluoreto de sodio
olafluor

valeratode estradiol
cloreto debenzalconio
acido ursodesoxicolico
lactogluconato de calcio
para-aminobenzoatode magnesio
hipromelose
complexo hidroxido ferrico-polimaltose
acido alendronico
cloromadinona
colquicina
zafirlucaste
dropropizina
policresaleno
alcool isopropilico
alcoolpropilico
etilsulfato demecetronio
acido nicotinico
etonogestrel0.015 mg/24 h
0.12 mg/24 hnuvaring1 unid;
acido espaglumico
clorodiazepoxido
brometo declidinio
butilescopolamina
pidolato de calcio
bicarbonatode sodio
imunoglobulina humana contra avaricela
imunoglobulina humana normal
cloreto de calcio
cloreto depotassio
cloreto de sodio
heparina sodica
oxerrutinas
lisado de colibacilos
benzilpenicilina benzatinica
benzilpenicilina potassica
cloreto de calcio
dexcetoprofeno
interferao alfa-2b
cetazolam
alcooldiclorobenzilico

acido alginico
acido estearico
erisimo officinalis
montelucaste
cloperastina
acido lactico
bromoprida
vacina contra a difteria, o tetano,a tosse convulsa e o haemophilustipo b
reviparina sodica
hidroxiquinolina
folitropina alfa
lutropina alfa
esteres etilicos 90 do acidoomega-3
vinburnina
aspartato de arginina
flubendazol
vaccinium myrtillus(antocianosidos)
cloreto de dequalinio
cinchocaina
omoconazol
toxina botulinica a
carbonatode magnesio
pirenoxina
tiropramida
vacina contra a hepatite a
vacina contra o papilomavirushumano
isobenzidrina
polistireno sulfonato de sodio
gonadotropina corionica
secnidazol
nicotinato de alfatocoferol
cassia angustifolia (fruto)
ispagula (mucilagem)
plantagoovata (sementes)
vacina viva contra o sarampo
cloreto de benzalconio
agua de hamamelis
aliscireno
vacina contra a gripe
vacina inactivada contra aencefalite provocada por picadade carraca

foliculina
carbonato de lantanio
hidroxiapatite
difluocortolona
vacina adsorvida pneumococicapoliosidica conjugada
complexo hidroxidoferrico-polimaltose
tansulosina
insulinaisofanica
parnaparina sodica
metamizol magnesico
interferao beta-1a
cobamamida
acetato de glatiramero
amissulprida
policresuleno
acido alendronico
valproato semisodico
micofenolato de mofetil
colagenase
acido quenodesoxicolico
alcatrao mineral
“heparinoide”
clorazepato dipotassico
aminaftona
calcitonina de salmao
acido para-aminossalicilico
tropissetrom
heparina sodica
salicilato dedietilamina
tiocolquicosido
imunoglobulina humana contra ahepatite b
altizida
clorofenoxamina
benzoato de benzilo
tinzaparina sodica
oleo essencial de hortela pimenta
cloreto de magnesio
cloretode potassio
levotiroxina sodica
enoxaparina sodica

acido dimecrotico
carbasalato calcico
salicilato demetilo
pepsinae outras associacoes
brometo de otilonio
senosido a
senosido b
octreotido
tianeptina
alfatocoferol
acidoascorbico
acido zoledronico
extracto de camomila
gluco-heptonatodecalcio
promestrieno
acido borico
oxido de zinco e outrasassociacoes
sene e outrasassociacoes
aspartatode potassio
cloreto de cetilpiridinio
mesoglicano sodico
cafeina
vacina viva contra a varicela
acido azelaico
loflazepato de etilo

A.4 UNC-Chapel Hill drug list [24]

Prescription medications that increase the risk of falls for patients 65 and older
Related UNC-Chapel Hill news release: http://uncnews.unc.edu/news/health-and-medicine/some-drugs-increase-risk-of-falling-unc-researchers.html

Generic Name (Brand Name)

Alprazolam (Xanax)
Amitriptyline (Elavil)
Amobarbital (Amytal)
Amoxapine (Asendin)
Aripiprazole (Abilify)
Baclofen (Lioresal)
Bupropion (Wellbutrin, Wellbutrin SR)
Buspirone (Buspar)
Butabarbital
Carbamazepine (Tegretol, Tegretol XR, Carbatrol)
Chloral hydrate
Chlorazepate (Tranxene)
Chlordiazepoxide (Librium, Limbitrol, Librax)
Chlorpromazine (Thorazine)
Citalopram (Celexa)
Clidinium-chlordiazepoxide (Librax)
Clomipramine (Anafranil)
Clonazepam (Klonopin)
Clozapine (Clozaril)
Codeine (Tylenol with Codeine)
Desipramine (Norpramin)
Diazepam (Valium)
Digoxin (Lanoxin)
Disopyramide (Norpace)
Divalproex sodium (Depakote, Depakote ER)
Doxepin (Sinequan, Zonalon, Prudoxin)
Duloxetine (Cymbalta)
Escitalopram (Lexapro)
Estazolam (Prosom)
Olanzapine (Zyprexa, Zyprexa Zydis)
Oxazepam (Serax)
Oxcarbazepine (Trileptal)
Oxycodone (Percocet)
Oxymorphone (Numorphan)
Paraldehyde (Paral)
Paroxetine (Paxil)
Pentobarbital (Nembutal)
Perphenazine (Trilafon)
Phenelzine (Nardil)
Phenobarbital
Phenytoin (Dilantin)
Pimozide (Orap)
Pregabalin (Lyrica)
Primidone (Mysoline)
Propoxyphene (Darvon, Darvocet)
Protriptyline (Vivactil)
Quazepam (Doral)
Ethosuximide (Zarontin)
Felbamate (Felbatol)
Fentanyl (Duragesic)
Fluoxetine (Prozac)
Fluphenazine (Permitil, Prolixin)
Flurazepam (Dalmane)
Fluvoxamine (Luvox)
Gabapentin (Neurontin)
Halazepam (Paxipam)
Haloperidol (Haldol)
Hydrocodone (Vicodin)
Hydromorphone (Dilaudid)
Imipramine (Tofranil)
Isocarboxazid (Marplan)
Levetiracetam (Keppra)
Levorphanol (Levo-Dromoran)
Lorazepam (Ativan)
Loxapine (Loxitane, Loxitane C)
Maprotiline (Ludiomil)
Mephobarbital
Meprobamate (Miltown, Equanil)
Mesoridazine (Serentil)
Methadone (Dolophine)
Methsuximide (Celontin)
Mirtazapine (Remeron)
Molindone (Moban)
Morphine (MS Contin)
Nefazodone (Serzone)
Quetiapine (Seroquel)
Risperidone (Risperdal)
Secobarbital (Seconal)
Sertraline (Zoloft)
Temazepam (Restoril)
Thioridazine (Mellaril)
Thiothixene (Navane)
Tiagabine (Gabatril)
Topiramate (Topamax)
Tranylcypromine (Parnate)
Trazodone (Desyrel)
Triazolam (Halcion)
Trifluoroperazine (Stelazine)
Trimipramine (Surmontil)
Venlafaxine (Effexor, Effexor XR)
Ziprasidone (Geodon)
Zolpidem (Ambien)
Zonisamide (Zonegran)

References

[Agi10] Aging Services of Minnesota. Aging Services of Minnesota. http://www.agingservicesmn.org/providers/clinical-excellence/fall-prevention-toolkit, 2010. [Online; accessed 3-Jan-2013].

[Agu12] Bruno Aguiar. Information Extraction from Medication Leaflets. Master's thesis, Faculdade de Engenharia da Universidade do Porto, 2012.

[AWA01] Cynthia L. Arfken, Jacquelyn Gardner Wilson, and Stephen M. Aronson. Retrospective Review of Selective Serotonin Reuptake Inhibitors and Falling in Older Nursing Home Residents. *International Psychogeriatrics*, 13(1):85–91, March 2001.

[CKG+08] Monica Campillos, Michael Kuhn, Anne-Claude Gavin, Lars Juhl Jensen, and Peer Bork. Drug target identification using side-effect similarity. *Science*, 321(5886):263–266, 2008. [Online; accessed 3-Jan-2013].

[CLC+11] Sung-Yen Chang, Chin-Feng Lai, Han-Chieh Josh Chao, Jong Hyuk Park, and Yueh-Min Huang. An Environmental-Adaptive Fall Detection System on Mobile Device. *Journal of Medical Systems*, 35(1):1299–1312, 2011.

[DBY+10] Jiangpeng Dai, Xiaole Bai, Zhimin Yang, Zhaohui Shen, and Dong Xuan. PerFallD: A pervasive fall detection system using mobile phones. *2010 8th IEEE International Conference on Pervasive Computing and Communications Workshops PERCOM Workshops*, 14(7):292–297, 2010.

[Dro00] Craig Drown. Alphabetical Listing of ATC drugs & codes . http://msupply.org.nz/files/atc_alphabetical.pdf, 2000. [Online; accessed 3-Jan-2013].

[Dye07] Owen Dyer. Simple precautionary measures can reduce numbers of falls in hospital. *BMJ British Medical Journal*, 334(7591):447, 2007.

[EBM02] Kristine E Ensrud, Terri L Blackwell, and Carol M Mangione. Central nervous system–active medications and risk for falls in older women. *Journal of the American Geriatrics Society*, 50(10):1629–1637, 2002.

[Etc08] David Etchison. Prescription medications that increase the risk of falls for patients 65 and older. uncnews.unc.edu/content/view/1364/107/, July 2008. [Online; accessed 3-Jan-2013].

[FPC05] FPCE. Fall Prevention Center of Excellence. http://www.stopfalls.org/, 2005. [Online; accessed 3-Jan-2013].

REFERENCES

[FTW+01] L P Fried, C M Tangen, J Walston, A B Newman, C Hirsch, J Gottdiener, T Seeman, R Tracy, W J Kop, G Burke, and M A McBurnie. Frailty in older adults: evidence for a phenotype. *The journals of gerontology Series A Biological sciences and medical sciences*, 56(3):M146–M156, 2001.

[Ga11] Vânia Guimarães. Phone Based Fall Risk Prediction. Master's thesis, Faculdade de Engenharia da Universidade do Porto, 2011.

[Gar10] Maria Giné Garriga. *Functional Performance in Physically Frail Community Dwelling Older Adults*. PhD thesis, Universitat Ramon Llull, 2010.

[GHG+10] Jonathan Gribbin, Richard Hubbard, John R F Gladman, Chris Smith, and Sarah Lewis. Risk of falls associated with antihypertensive medication: population-based case-control study. *Age and ageing*, 39(5):592–7, September 2010.

[Har81] D Harman. The aging process. *Proceedings of the National Academy of Sciences of the United States of America*, 78(11):7124–7128, 1981. [Online; accessed 3-Jan-2013].

[Hee04] W.J. Heeringa. *Measuring Dialect Pronunciation Differences Using Levenshtein Distance*. Groningen dissertations in linguistics. 2004.

[HLL07] Sirpa Hartikainen, Eija Lönnroos, and Kirsti Louhivuori. Medication as a risk factor for falls: critical systematic review. *The journals of gerontology Series A Biological sciences and medical sciences*, 62(10):1172–1181, 2007.

[Hos09] Mansfield District Hospital. Falls related medication side effects. 2009. Last Checked: 13/01/2013.

[Hub03] R. Hubbard. Exposure to Tricyclic and Selective Serotonin Reuptake Inhibitor Antidepressants and the Risk of Hip Fracture. *American Journal of Epidemiology*, 158(1):77–84, July 2003.

[ICE04] ICECI. International Classification of External Causes of Injuries (ICECI) July 2004 A Related Classification in the World Health Organization Family of International Classifications. page 245, July 2004.

[KD04] G Kondrak and B Dorr. Identification of confusable drug names: a new approach and evaluation methodology. In *Proceedings of the 20th international conference on Computational Linguistics*, COLING '04, Stroudsburg, PA, USA, 2004/// 2004. Association for Computational Linguistics, Association for Computational Linguistics.

[KP09] Hamed Ketabdar and Tim Polzehl. Fall and emergency detection with mobile phones. In *Proceeding of the eleventh international ACM SIGACCESS conference on Computers and accessibility ASSETS 09*, page 241. ACM SIGACCESS, ACM Press, 2009. [Online; accessed 3-Jan-2013].

[KPY+03] Karen D Kelly, Will Pickett, N Yiannakoulias, Brian H Rowe, Don P Schopflocher, Larry Svenson, Don C Voaklander, and Kaven D Kelly. Medication use and falls in community-dwelling older persons. *Age and ageing*, 32(5):503–9, September 2003.

REFERENCES

[LAM+98] Barbara Liu, Geoff Anderson, Nicole Mittmann, Teresa To, Tami Axcell, and Neil Shear. Use of selective serotonin-reuptake inhibitors or tricyclic antidepressants and risk of hip fractures in elderly people. *The Lancet*, 351(9112):1303 – 1307, 1998. Last Checked: 13/01/2013.

[LPE03] Debbie A Lawlor, Rita Patel, and Saha Ebrahim. Association between falls in elderly women and chronic diseases and drug use: cross sectional study. *British Medical Journal*, 327(7417):712–717, 9 2003.

[MB00] J. Mintzer and A. Burns. Anticholinergic side-effects of drugs in elderly people. *Journal of the Royal Society of Medicine*, 93(9):457–462, September 2000. [Online; accessed 3-Jan-2013].

[MGMO12] S.W. Muir, K. Gopaul, and M.M. Montero Odasso. The role of cognitive impairment in fall risk among older adults: a systematic review and meta-analysis. *Age Ageing*, 41(3):299–308, 2012.

[MGMS80] M. Patricia Murray, Gena M. Gardner, Louise A. Mollinger, and Susan B. Sepic. Strength of isometric and isokinetic contractions: Knee muscles of men aged 20 to 86. *Physical Therapy*, 60(4):412–419, April 1980. Last Checked: 13/01/2013.

[MH11] Infarmed Ministry and Portuguese Health. Prontuario Terapêutico Online. http://www.infarmed.pt/portal/page/portal/INFARMED/PUBLICACOES/PRONTUARIO, 2011. [Online; accessed 3-Jan-2013].

[MM97] C Mustard and T Mayer. Case-control study of exposure to medication and the risk of injurious falls requiring hospitalization among nursing home residents. *American journal of epidemiology*, 145(8):738–45, April 1997.

[Mor09] Janice M. Morse. *Preventing Patient Falls*. Springer Publishing Company, second edi edition, 2009.

[NHMP96] C I Neutel, J P Hirdes, C J Maxwell, and S B Patten. New evidence on benzodiazepine use and falls: the time factor. *Age and ageing*, 25(4):273–8, July 1996.

[ORM09] Pamela L Owens, C Allison Russo, and Ryan Mutter. STATISTICAL BRIEF # 80 Emergency Department Visits for Injurious Falls among the Elderly. *Injury Prevention*, pages 1–11, 2009.

[Phy00] Am Fam Physician. What Causes Falls in the Elderly? How Can I Prevent a Fall? *American Family Physician*, 61, 2000.

[PMT+01] C Pierfitte, G Macouillard, M Thicoïpe, a Chaslerie, F Pehourcq, M Aïssou, B Martinez, R Lagnaoui, a Fourrier, B Bégaud, J Dangoumau, and N Moore. Benzodiazepines and hip fractures in elderly people: case-control study. *BMJ (Clinical research ed.)*, 322(7288):704–8, March 2001.

[PVR+00] Angelina Passaro, Stefano Volpato, Franco Romagnoni, Nadia Manzoli, Giovanni Zuliani, Renato Fellin, and Gruppo Italiano. Benzodiazepines with different half-life and falling in a hospitalized population : the GIFA study. 53:1222–1229, 2000.

[RBBL03] E Z Riefkohl, HL Bieber, MB Burlingame, and DT Lowenthal. Medications and falls in the elderly: a review of the evidence and practical considerations. *PharmacyTherapeutics*, 28(11):724–733, 2003.

REFERENCES

[RCV03] Suely Rozenfeld, Luiz Antonio Bastos Camacho, and Peixoto Veras. Medication as a risk factor for falls in older women in Brazil. *Revista panamericana de salud pública = Pan American journal of public health*, 13(6):369–375, June 2003. [Online; accessed 3-Jan-2013].

[RJR94] L Z Rubenstein, K R Josephson, and a S Robbins. Falls in the nursing home. *Annals of internal medicine*, 121(6):442–51, September 1994.

[RL93] Robin Ruthazer and A L Lewis. Elderly People in Long-Term Care. 83(5), 1993.

[SBMdPS10] Isabel Segura-Bedmar, Paloma Martínez, and César de Pablo-Sánchez. Extracting drug-drug interactions from biomedical texts. *BMC Bioinformatics*, 11(5):P9, 2010.

[SH08] Human Services and Ronda G Hughes. Patient Safety and Quality : An Evidence-Based Handbook for Nurses. *Health San Francisco*, 90(8):601–602, 2008.

[Sow00] John F Sowa. *Knowledge Representation: Logical, Philosophical, and Computational Foundations*, volume 27 of *Computer Science Series*. Brooks/Cole, 2000. [Online; accessed 3-Jan-2013].

[SPA09] SPANTEC GMBH. Fall Detector. https://play.google.com/store/apps/details?id=at.spantec.securemotion.falldetector&feature=search_result#?t=W251bGwsMSwxLDEsImF0LnNwYW50ZWMuc2VjdXJlbW90aW9uLmZhbGxkZXRlY3RvciJd, 2009. [Android App; accessed 3-Jan-2013].

[SPP10] Vick Scott, Marty Pearce, and Cate Pengelly. Deaths Due to Falls Among Canadians Age 65 and Over. Technical report, Public Health Agency of Canada, 2010.

[SSvD+08] Alice C Scheffer, Marieke J Schuurmans, Nynke van Dijk, Truus van der Hooft, and Sophia E de Rooij. Fear of falling: measurement strategy, prevalence, risk factors and consequences among older persons. *Age and ageing*, 37(1):19–24, January 2008.

[TMC11] Carlo Tacconi, Sabato Mellone, and Lorenzo Chiari. Smartphone-Based Applications for Investigating Falls and Mobility. *Proceedings of the 5th International ICST Conference on Pervasive Computing Technologies for Healthcare*, pages 258–261, 2011.

[Ton10] Lorena Tonarelli. Consequences of Falls in the Elderly Falling Increases the Risk of Disability and Mortality in Older Age. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1298101/, 2010. [Online; accessed 3-Jan-2013].

[UA11] Health & Human Services U.S. National Library of Medicine, National Institutes of Health and Freedom of Information Act. Unified Medical Language System. http://www.nlm.nih.gov/research/umls/, 2011. [Online; accessed 3-Jan-2013].

[WHO13] WHO. World Health Organisation. http://gamapserver.who.int/gho/interactive_charts/mbd/life_expectancy/atlas.html, 2013. [Online; accessed 3-Jan-2013].

[WId12] DIMDI WIdO. Anatomisch-therapeutisch- chemische Klassifikation mit Tagesdosen. http://wido.de/fileadmin/wido/downloads/pdf_arzneimittel/atc/wido_arz_amtl_atc-index_0111.pdf, 2012. [Online; accessed 3-Jan-2013].

[ZDH+06] G Ziere, J P Dieleman, a Hofman, H a P Pols, T J M van der Cammen, and B H Ch Stricker. Polypharmacy and falls in the middle age and elderly population. *British journal of clinical pharmacology*, 61(2):218–23, February 2006.